Sentence-final particle vs. sentence-final emoji: The syntax-pragmatics interface in the era of CMC

Chenchen Song*

Zhejiang University

Abstract

In this article, I present a formal linguistic analysis of affective emojis (i.e., emojis that are used to add tones to text messages) in computer-mediated communication (CMC) and lay out some preliminary thoughts on CMC linguistics. My analysis, which builds on the root-based approach to semilexical elements in generative syntax, separates CMC data with affective emojis into a non-CMC-specific part (i.e., the linguistic text) and a CMC-specific part (i.e., the emoji), with the latter functionally wrapping around the former and thereby setting its tone. This analysis can be applied to other CMC-specific affective elements too, such as memes and background music. The special nature of the digital modality has nontrivial ramifications for CMC linguistics. I argue that until the "legibility conditions" of the cyber-digital system are ascertained, the safest linguistic tools to use in research on CMC-specific phenomena are those that are not designed exclusively for the cognitive domain of language.

Keywords: emoji, computer-mediated communication, syntax, pragmatics, sentence-final particle, modal particle

1 Introduction[1]

Haralambous (2020:12) introduces grapholinguistics as "the discipline dealing with the study of the written modality of language" and points out that the reason why it has received little recognition is because writing has long been viewed "just as an accidental secondary representation of language." This position dates back to at least Ferdinand de Saussure's *Course in General Linguistics* (originally published in 1916):

*Thanks to the audience at the 2022 Grapholinguistics in the 21st Century Conference (6/8/2022) and the audience at the Cambridge SyntaxLab (6/28/2022) for constructive feedback. Thanks to Xiaoke Bu, Chunan Li, Shangze Li, Li Nguyen, Michele Sanguanini, Ke Wu, and Ruikang Zhang for participating in my survey.

[1]Abbreviations: AP = affective punctuation, C-D = cyber-digital, CL = classifier, CMC = computer-mediated communication, Conj = conjunction, CP = complementizer phrase, CRS = currently relevant state, DECL = declarative, DISP = disposal, DP = determiner phrase, EMPH = emphasis, EP = emotion phrase, MP = modal particle, NP = noun phrase, NumP = number phrase, PL = plural, POSS = possessive, Q = question marker, REL = relative clause marker, SFE = sentence-final emoji, SFP = sentence-final particle, TP = tense phrase, TU = text unit, v^*P = transitive light verb phrase, VP = verb phrase

> Language and writing are two distinct systems of signs; the second exists for the sole purpose of representing the first. The linguistic object is not both the written and the spoken forms of words; the spoken forms alone constitute the object. (Saussure 2011)

I agree with Saussure. It is a basic fact that human language, either spoken or signed, does not depend on writing. That said, however, I wonder whether Saussure would still have put his view in such an absolute tone if he had had the chance to time-travel to the 2020s and see how human beings are staying in touch nowadays.

Face-to-face (or voice-to-voice) communication is certainly still with us, but in the meantime, modern technology has made computers, smartphones, and the like an indispensable additional channel of communication. Given this revolutionary change of lifestyle, it is unclear to me to what extent we can confidently assert that writing—or really typing (e.g., texting, tweeting)—is still strictly secondary to oral language. Among others, many CMC-specific communicative elements—such as emojis, memes, and GIFs—have never existed in oral speech and never will. They are native to the digital modality of communication instead. In this article, I present an emoji-centered case study of CMC and hope to convince readers that we need to rethink the relation between language and writing/typing in the 21st century.

Emojis play an increasingly important role in our day-to-day lives, in that they compensate for the lack of nonverbal or "paralinguistic" (Carey's 1980 term) cues in online textual communication. As suggested by Gawne and McCulloch (2019), the place of emojis in computer-mediated communication (CMC) is equivalent to that of "tone of the voice and body language in face-to-face communication." It is fair to say that emojis are becoming an integral part of human language in the digital age. As a linguist, I am most interested in the following questions:

1. What is the cognitive nature of CMC data involving emojis? Is the normal tool kit from linguistics sufficient for an adequate analysis of them?

2. If it turns out that the nature of CMC data is fundamentally different in certain aspects from that of conventional linguistic data, then which part of the linguistic tool kit is still applicable to their analysis?

The rationale behind these questions is as follows. Modern linguistics, in particular its generative branch (Chomsky 1957 et seq.), is established on the hypothesis that our language capacity is supported by a dedicated mental organ—the language faculty. This is a computational system that generates complex structures out of basic linguistic units (e.g., words). The language faculty interfaces with two other cognitive systems: the sensorimotor system and the conceptual-intentional system (Chomsky 1995). The former is where abstract linguistic structures get externalized as physical signals, and the latter is where they get interpreted as language-based thoughts. A major goal of contemporary theoretical linguistics is to specify how information flows from the computational system to the interface systems. For instance, linguists have proposed many operations in the past few decades to tackle the question of how hierarchical syntactic structures are converted to linear strings usable in the oral-auditory modality (see Biberauer and Roberts 2013 for an impression of the complexity of this issue). Due to the central status of linearization in pre-CMC-era linguistics, quite a few theoretical tools initially designed for linearization purposes alone have subsequently been made part of the core design of the language

faculty (such as "cyclic spell-out" and its latest incarnation Phase Theory; Chomsky 2001).

My questions above are based on the concern that, if CMC is not confined by the naturally evolved communicative modalities (including but not limited to the oral-auditory modality) or their requirements, then what theoretical linguistic tools can we still apply to CMC data, and what tools must we refrain from using? These are big questions whose settling calls for much more research and community efforts. For the limited purpose of this article, I wish to demonstrate the applicability of just one formal linguistic tool: root categorization.

As has been mentioned, my case study is centered on emojis. In particular, the emoji usage described above is *affective* in nature. Affective emojis convey speaker attitudes or tones. Emojis can also be used in a *nonaffective* way. This is the situation where an emoji is simply used as an icon for a verbal concept, usually directly substituting for a word. See (1) for an illustration.[2]

(1) a. Great idea 👍 I'm in 😊

b. If I were in Detroit, I'd give you a 🎁 (adapted from Maier 2021:4)

The two emojis in (1a) are used affectively. They respectively express an approving tone and a genuinely happy tone. By contrast, the emoji in (1b) is used nonaffectively. It merely represents a gift and can be directly replaced by the word "gift." The two types of emoji usage above may be alternatively described as use-conventional vs. truth-conditional or non-at-issue vs. at-issue (Grosz et al. 2021, Maier 2021, Pierini 2021), the latter being based on a piece of terminology in Potts (2005). In what follows, I will stick to the affective vs. nonaffective terminology.

I focus on affective emojis in this article. Note that the two affective emojis in (1a) are both attached to the end of the sentence they accompany—or more exactly the *text unit*, since "Great idea" is not a complete sentence. This syntactic property is true of affective emojis in general. Hence, I also call affective emojis *sentence-final emojis* (SFEs). I choose this designation because the above combination of syntactic and semantic properties—namely, being sentence-final and expressing speaker affects—is reminiscent of a class of vocabulary elements in oral languages, especially in East and Southeast Asian languages, which have been called "sentence-final particles" (SFPs) in the linguistic literature (see, e.g., Cheng and Tang 2022 and Morita 2018). See (2) for some examples from Mandarin Chinese, which is also my main source of data.[3]

(2) a. *xià xuě le* ***ye*** [Mandarin Chinese]
fall snow CRS SFP
'It snowed. (excited tone)'

b. *xià xuě le* ***a***
fall snow CRS SFP
'It snowed. (surprised tone)'

[2]I generally use Apple emojis in this article but will switch to alternative versions in cited examples, since different implementations of the same emoji often have subtle differences in the exact affects they convey (see §2.3).

[3]I follow the standard practice in linguistics and present non-English examples in a three-line format: the first line is the original example (or its romanization, if the original language has a non-Latin script), the second line is a verbatim glossing of the example (in an English-based metalanguage), and the third line is a more natural English translation.

c. *xià xuě le* ***you***
fall snow CRS SFP
'It snowed. (kind reminder tone)'

d. *xià xuě le* ***ha***
fall snow CRS SFP
'It snowed. (harmony-seeking tone)'

In (2), the same new situation "it is snowing (now)" is reported in four different tones, which are encoded in four different SFPs. In CMC, the same communicative effects can be achieved via affective emojis, as in (3).

(3) *xià xuě le* 😄/😮/🙂/😅 [Mandarin Chinese]
fall snow CRS SFE
'It snowed. (excited/surprised/reminder/harmony-seeking tone)'

The particle-emoji parallelism above is striking. One may even conclude that SFEs are the digital counterpart of SFPs. Indeed, the two types of affect-expressing elements have been given a unified linguistic analysis in Song (2019). However, in this article I will show that despite their functional similarity, we cannot put SFPs and SFEs in the same category. While the former are an integral part of oral speech, the latter are first-class citizens of CMC (and CMC alone). I will present three arguments that bear out the categorial distinction between SFPs and SFEs:

1. SFPs and SFEs can and often do co-occur.

2. SFPs are a closed class, whereas SFEs are an open class.

3. The positioning of affective emojis is not influenced by crosslinguistic word order variation, whereas that of affective particles is.

The three arguments will be elaborated one by one. After that, I will propose a new linguistic analysis for SFEs, which is based on the Generalized Root Syntax theory in Song (2019).

The rest of this article is structured as follows. In Section 2, I present my arguments against an identical linguistic treatment of SFPs and SFEs. In Section 3, I present my new analysis of affective emojis. In Section 4, I discuss the implication of my case study for the field of CMC linguistics in general.Section 5 concludes.

2 SFP and SFE are different categories

In this section, I comparatively examine the linguistic behavior of SFPs and that of SFEs and argue that they should not be treated as the same category. I begin with a note on SFP taxonomy (§2.1), then move on to present my three arguments (§2.2–2.4), and finally make a digression on sentence-initial emojis (§2.5), showing that they are not counterexamples to my generalization. I end the section with an interim summary (§2.6) that prepares the ground for my theoretical analysis.

Table 1: A taxonomy of Mandarin Chinese SFPs (adapted from Paul 2014)

Type	Characterization	Examples
I	Tense/Aspect	*le* 'currently relevant state' *láizhe* 'recent past' ne_1 'continued state'
II	Sentence type	*ma* 'yes-no question' *ba* 'imperative' ne_2 'follow-up question'
III	Attitude	*o* 'mild reminder tone' *a*/*ya* 'surprised tone' ne_3 'exaggerating tone'

2.1 SFP taxonomy

SFPs are not a homogeneous category. According to Paul (2014), the SFPs in Mandarin Chinese fall in three types, as shown in Table 1.

Type I SFPs in Mandarin are tense or aspect markers, such as the currently relevant state marker *le*, the effect of which partly overlaps with that of the perfect in English. Thus, in the "snowing" examples in (2), a more accurate paraphrase of the statement "it snowed" is "it has snowed some time ago, and that state of affairs is relevant to the current situation we are in (e.g., there is snow on the ground)." Type II SFPs are sentence type markers, such as the yes-no question marker *ma*, which turns a proposition into a yes-no question and is similar in effect to French *est-ce que*. Thus, while *xià xuě le* 'it snowed' is a statement, *xià xuě le ma* 'It snowed?' is a question. Type III SFPs are attitude markers. All four examples in (2) are of this type. This is also the type of SFP that I focus on in this article. Hereafter, by "sentence-final particle" I only mean Type III SFPs.

2.2 Argument I: SFPs and SFEs can co-occur

The first reason why SFPs and SFEs should not be treated as the same category is that they can and often do co-occur in the same sentence. For instance, the patterns in (2) and (3) can be combined into (4).

(4) a. *xià xuě le* ***ye*** 😄 [Mandarin Chinese]
fall snow CRS SFP SFE
'It snowed. (excited tone)'

b. *xià xuě le* ***a*** 😮
fall snow CRS SFP SFE
'It snowed. (surprised tone)'

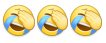
c. *xià xuě le* ***you*** 🙂
fall snow CRS SFP SFE
'It snowed. (kind reminder tone)'

d. *xià xuě le* ***ha*** 😅
fall snow CRS SFP SFE
'It snowed. (harmony-seeking tone)'

In fact, the forms in (4) are more natural than those in (3), because the retention of the SFPs makes the messages more speech-like, while the addition of the SFEs helps further highlight the tones in the SFPs. Such SFP-SFE co-occurrence is common in CMC data. See (5) for more examples from the social media website Sina Weibo (henceforth Weibo), which is the Chinese equivalent of Twitter.

(5) a. *wǒ měitiān dōu zài zhíbō* ***o*** *qīn*  [Mandarin Chinese]
I everyday all be.at live-stream SFP dear SFE
'For your information, dear, I'm live-streaming everyday. (teasing tone)'

b. *nǐ de wǎng-míng hěn fúhé nǐ* ***o*** 
you POSS Internet-name very suit you SFP SFE
'Just saying, your profile name fits you very well. (jocularly cheeky tone)'

c. *wǒ zěnme jìde hǎoxiàng shì liú bǎ tā chuài le* ***a*** 
I how remember likely is Liu DISP her dump CRS SFP SFE
'How come I vaguely remember that it was Liu who had dumped her? (highly amused tone)'

d. *nǚ míngxīng shēngrì kuàilè* ***o*** 
female star birthday happy SFP SFE
'Superstar girl, happy birthday! (cute fangirl tone)' (Weibo)

Like many Asian social media platforms, Weibo has its own emojis, which are outside the Unicode list. Nevertheless, the usage of the Weibo-specific emojis in (5) is not different from that of the Unicode emojis we have seen. Moreover, in these examples, the SFEs are not translations of the SFPs. Rather, in each example, the affects in the SFP and the SFE combine into a new and more subtle tone. I will come back to platform-specific, non-Unicode emojis in Section 2.3. Specifically, (5a), (5b), and (5d) share the same base tone—the mild reminder tone encoded in the SFP *o*—which is further forged by the additional SFEs in three different molds, respectively into a teasing reminder, a jocularly cheeky reminder, and a fangirlish reminder. Similarly, the surprised-tone SFP *a* in (5c) combines with the "allow me to do a sad face" emoji (repeated three times) to yield a seemingly surprised but actually highly amused tone.

The productive co-occurrence of SFPs and SFEs is a clear indication that the two types of affective element instantiate different linguistic categories, with a category being understood as an equivalence class in terms of linguistic behavior. To begin with, linguistic elements of the same category are usually in complementary distribution, which is partly what motivates linguists to define them as a category in the first place. See (6) for two familiar examples. The asterisk indicates that the expression after it is ill-formed.

(6) a. this book, that book, *this that book (Demonstrative)

b. I like reading, you like reading, *I you like reading (Pronoun)

This and *that* are in the same category (Demonstrative) because they can freely substitute for each other without affecting grammaticality and cannot be used simultaneously, and similarly for the nominative pronouns *I* and *you*. Note that the conception of category adopted here is a fine-grained one. Supposing categories are hierarchically organized in their ontology into super- and subcategories, I only consider elements of the same smallest subcategory as categorially equivalent. Thus, while nonnominative pronouns like *me* and

him are also in the general category Pronoun, they are not equivalent to nominative pronouns.

Furthermore, when SFPs and SFEs co-occur, their order cannot be switched. That is, the SFP slot can only be filled by oral-language particles, while the SFE slot can only be filled by emojis (or other similar digital symbols, such as emoticons). Sentences like the following are unacceptable.

(7) a. **xià xuě le* 😄 ***ye*** [Mandarin Chinese]
fall snow CRS SFE SFP
'It snowed. (excited tone)'

b. **nǐ de wǎng-míng hěn fúhé nǐ* 😁 ***o***
you POSS Internet-name very suit you SFE SFP
'Just saying, your profile name fits you very well. (jocularly cheeky tone)'

This restriction is unexpected if the two types of affective elements are categorially equivalent.

2.3 Argument II: SFEs are an open class

The second reason why SFPs and SFEs should not be treated as the same category is that SFPs are a closed class, while SFEs are an open class. Thus, even if they were in the same category, that category would still be a hybrid one, with two heterogeneous subcategories, which brings us back to the ontological issue mentioned above.

The inventory of SFPs in Sinitic languages is not particularly small, especially if we take all three subtypes in Section 2.1 into consideration. However, they are still a closed class, which means that the set of SFPs in a Sinitic language is stably fixed in an extended period of time. Take Mandarin Chinese for example. Although scholars hold varied opinions on its number of SFPs, that number is generally assumed to be under 30. Among others, Chao (1968) lists 26 (including many borderline items), Sun (1999) lists 28 (for all Mandarin subvarieties throughout the 19th and 20th centuries), and Li and Thompson (1981) list 6 (only the most common ones).

By contrast, the inventory of SFEs is much larger and also keeps expanding. This is evidenced by four observations:

1. New face emojis are created every year.
2. Nonface emojis can be used affectively too.
3. There are plenty of platform-specific, non-Unicode affective emojis.
4. There are various quasi emojis (e.g., emoticons, affective punctuation marks).

In what follows, I will elaborate on these observations one by one. First, new face emojis are being regularly created, almost on a yearly basis. See (8) for some examples.

(8) 2018: 🥰, 🥵, 🥶, 🥳, 🥴, 🥺
2019: 🤭
2020: 🥲, 🤓, 😶‍🌫️, 😮‍💨, 😵‍💫
2021: 🫣, 🤫, 🫨, 🫠, 😐, 🫤, 🥹 (Emojipedia)

Face emojis are naturally affective, so their constant expansion is clear evidence for the open-class nature of SFEs. However, not all affective emojis are face emojis, and that brings us to the second piece of evidence for the open-class nature of SFEs—namely, that nonface emojis can also be used affectively. When studying affective emojis, we should not limit our attention to just face-based ones (pace Grosz et al. 2021).

The affective use of nonface emojis is highly versatile. Some more systematic ones are hand emojis like 🤝, 🙏, and 👏 and heart emojis like ❤️, 💔, and 💝. There are also less systematic ones, such as those in (9). For authenticity's sake, I have retained the spelling and emoji style (i.e., the Twitter version) of the original tweets.

(9) a. Perfect art! So talented artist 🔥

b. had 'hug' been a little more second longer, she would've elbowed one of these queens out. just saying 💅

c. Every woman wants a man who's hard-working and ambitious until it's the weekend and he plans on working 🐸☕ (Twitter)

In all these examples, the nonface emojis are clearly used affectively, in that they serve to convey speaker attitudes. The fire emoji in (9a) conveys an enthusiastically admiring tone, the nail polish emoji in (9b) conveys a nonchalant tone, and the frog-and-hot-beverage emoji compound in (9c) conveys a sarcastic tone.[4] While the above affective uses are all largely conventionalized—in that they are regularly used in the relevant affective senses—there are also more ad hoc affective uses of nonface emojis. The temporary nature of such usage is especially clear in cases where a sequence of emojis are randomly put together to convey a strong emotion, as exemplified in (10).

(10) a. Awesome cooperation🤜👏🔥

b. Can't wait to see your performance, I hope everything goes well. please stay healthy guys💛🐥🌻

c. my baby he looks so handsome and cute🫶💞🧸❤️‍🔥🍬💝🍭❤️ (Twitter)

The three tones added by the emoji sequences in (10) are respectively strongly celebratory, highly affectionate and caring, and extremely cloying. The third example is particularly interesting, as the three non-heart-based emojis in the sequence (i.e., 🧸, 🍬, and 🍭) further forge the strongly loving tone in the heart emojis into a cute-baby-loving tone, even though the "baby" in the sentence does not refer to a real infant (but refers to a grown-up man instead).

The third piece of evidence for the open-class status of SFEs is the abundance of platform-specific affective emojis, which are often outside the Unicode list. Asian social media platforms are particularly creative in this respect, where many emojis do not have counterparts on Western platforms. We have seen a few examples in (5). Below are more examples from three popular Chinese platforms, which respectively correspond to Twitter, Facebook/WhatsApp, and TikTok in the West.

[4]While the tones in 🔥 and 💅 are stably fixed, the tone in 🐸☕ is less so. According to Emojipedia, this combination could be used for gossip or sarcasm or be associated with trolling or the alt-right.

(11) a.  (Weibo)

b.  (WeChat)

c.  (Douyin)

Some of the emojis in (11) make social-cultural references and so cannot be immediately understood by non-Chinese-speakers. Take the watermelon-eating emoji for instance, which has three slightly different implementations on Weibo (), WeChat (), and Douyin () and sometimes is just simplified as  . The affective use of this emoji makes reference to the slang expression in (12).

(12) *bù-míng zhēnxiàng de chī-guā qúnzhòng* [Mandarin Chinese]
not-understand truth REL eat-melon masses
'people who are merely watching an event from afar while eating watermelon but do not understand what is really going on'

The phrase is often shortened as *chī-guā qúnzhòng* 'watermelon-eating masses' and has given rise to a derivative *chī-guā* 'to be a member of the watermelon-eating masses'. The corresponding emoji is usually used to convey a rubbernecking onlooker's attitude, as in (13).

(13) a. *xiànzài bàochū shénme xīnwén wǒ dōu bù xīqí le* 
now break whatever news I EMPH not curious CRS SFE
'Nowadays I'm no longer shocked by whatever news. (onlooker's tone)'

b. *wúcháng fēnxiǎng gěi nǐmen* 
gratis share to you all SFE
'I'm sharing these (celebrity scandals) with you for free. (onlooker's tone)' (Weibo)

Note that the three platform-specific implementations of the watermelon-eating emoji mentioned above are not completely equivalent. Intuitively, the Weibo version  has a more "none of my business" attitude, the WeChat version  has a more gossipy feeling, while the Douyin version  has a more "peanut gallery" effect. Such subtle tonal variation in a sense makes the inventory of SFEs even larger—because if two variants of the same emoji convey different tones, then they may as well be treated as different emojis.

Internet users' intuition of the tonal variation in affective emojis is impressively nuanced. I conducted a small-scale survey on whether the platform-specific implementations of the eye-rolling emoji convey the same tone, and the general answer I got was No. See Table 2 for the detailed responses. Note that the two QQ[5] versions are both animated, but I can only present them as static screenshots here. To further illustrate the rich intuition Internet users possess about emoji usage, I quote the following additional comments from my respondents:

> Compared with the other eye-rolling emojis, this animated one [QQ 2] ... adds extra absurdity and humor. With the smiling, there is also a slightly sarcastic tone. I think it is a mixture of complex emotions and subtle feelings. Thus, personally, I find it peculiarly lovely. (User 5)

[5]QQ is a Chinese instant messaging software service.

For me, emojis with a nonflat mouth are more negative than emojis with a flat mouth, which are in turn more negative than emojis with an open mouth. So, here the Twitter version of the eye-rolling is more negative than the Apple version, which in turn is more negative than the first WeChat version. The second QQ version is different from all the others. I tend to express the emotion of sarcasm or fake politeness when using it. (User 7)[6]

[6]Since User 7 only provided this general remark, I did not include their response in Table 2.

Table 2: Survey results on the tonal differences between platform-specific eye-rolling emojis

	🙄 Apple	😕 Twitter	😬 WeChat 1	🤨 WeChat 2	🙄 Weibo	🙄 QQ 1	🫤 QQ 2
User 1	"I can't even," jaded	disappointed	"eye-avoidance," embarrassed	disappointed	disappointed & sad	slightly embarrassed or a bit cheeky	amused (for chaos or minor confusion)
User 2	slightly annoyed	a bit sad	wondering	confused	slightly indifferent or skeptical	slightly naughty	silly
User 3	speechless (negative)	negative attitude	playing innocent, "not me not me"	pretending to be angry	negative attitude	playing innocent, "not me not me"	speechless (negative)
User 4	speechless	speechless & unhappy	"I don't wanna hear"	pretending to be angry	speechless (friendlier)	"I don't wanna hear" (cuter)	totally speechless, "death smile"
User 5	real eye-rolling (highly negative)	≈Weibo	≈QQ 1	pretending to be angry	a bit of disdain	a bit shocked	humorously sarcastic
User 6	real eye-rolling	confused	pretending to be confused	arrogant	pondering	pretending to be confused	backhanded compliment

Finally, the abundance of affective quasi emojis, such as emoticons and affective punctuation marks, is also evidence that SFEs are an open class. Modern-day emoticons are far more versatile than sideways smileys like :-) and :D. Once again, Asian Internet users are particularly creative in this realm. See (14) for some examples of Japanese kaomojis.[7]

(14)

(^◡^)	(*￣▽￣)b	(.•◡•.)	(〃^▽^〃)	o(•◡•)o
(o˘◡˘o)	o(≧▽≦)o	(´｡•ᵕ•｡`)	\(★ω★)/	\(≧▽≦)/
ヽ(o^▽^o)ノ	ヽ(o^ ^o)/	(*^‿^*)	╰(￣∀￣)╯	(≧◡≦)
(((o(*°▽°*)o)))	(o´▽`o)	(@^◡^)	(o･ω･o)	(￣ω￣)
(´｡• ω •｡`)	ヽ(・∀・)/	｡.:☆*:･'(*⌒―⌒*)))	<(￣‿￣)>	(o^▽^o)

Kaomojis are highly popular in Japan and China—so much so that smartphone keyboards now have a special section for them. In addition, Chinese and Korean speakers sometimes use special emoticons made up of Chinese/Korean characters, such as 囧 (an embarrassed face) in Chinese[8] and ㅇㅁㅇ (a shocked face) in Korean.

As for affective punctuation marks, apart from the conventional question and exclamation marks (and their various combinations), the ellipsis and the tilde are good examples too. The former is popular around the world and usually signals hesitation or silence, while the latter is mainly popular in Asia and signals cuteness or a softened tone. See (15) for an illustration.[9]

(15) a. *zhēnde ma ...* [Mandarin Chinese]
real Q AP
'Really? (hesitant tone)'

b. *bāng wǒ mǎi dōngxi* ～～～
help me buy stuff AP
'Help me buy something please. (cute tone)'

The ellipsis in (15a) conveys hesitation, which may be translated as "alright" or "whatever" depending on the context. The tildes in (15b), on the other hand, create a friendly and cute-sounding effect, which is important in texting, since otherwise the message sounds rather abrupt.

2.4 Argument III: Affective emoji positioning is not influenced by word-order variation

My third argument for the categorial difference between SFPs and SFEs is based on a more general observation about affective particles. SFPs in Chinese and other Asian languages are a major type of affective particle in human language, but they are not the only type. Among others, the modal particles in German (and some other Germanic languages, such as Dutch; see, e.g., Fehringer and Cornips 2019) are also affective, in that they also serve to convey speaker tones or attitudes. See Table 3 for a selection of common German modal particles and see (16) for some concrete examples.

[7] These examples are extracted from kaomoji.ru/en. (last visited on 10/22/2022)

[8] The Chinese character 囧 originally means "bright" and is pronounced *jiǒng*, but its usage as an emoticon has nothing to do with its original meaning and is merely a shape-based recycling.

[9] Chinese speakers often use a sequence of Chinese-style periods (◦◦◦) in place of ellipsis dots (...), and for some speakers the former conveys an even stronger hesitant tone, but I abstract away from this subtlety here.

Table 3: Some German modal particles (Durrell 2021:§9.1)

Particle	Connotation
halt	an attempt by the speaker to put an end to any discussion because the situation does not allow any alternatives
ja	appealing for agreement, expressing surprise, intensifying commands
mal	making the tone sound less blunt
doch	typically used to try to persuade the listener of the speaker's point of view, usually expressing a contradiction or disagreement
nun	signaling dissatisfaction with a previous answer or that the speaker considers the topic exhausted
eben	typically expressing a confirmation that something is the case

(16) a. *Das ist **halt** so.* [German]
that is MP so
'But there, that's how it is. (there's-nothing-one-can-do tone)'

b. *Ihr habt **ja** früher zwei Autos gehabt.*
you.PL have MP earlier two cars had
'Of course, you used to have two cars. (as-we-all-know tone)'

c. *Ich kann ihn nicht überreden. Er ist **eben** hartnäckig.*
I can him not convince he is MP obstinate
'I can't convince him. He's just obstinate. (it-can't-be-helped tone)'
(Durrell 2021:§9.1)

As we can see, the position of affective modal particles in German is consistently sentence-medial instead of sentence-final. This shows that the syntactic position of affective particles, like that of most other elements of oral languages, is subject to crosslinguistic variation. By contrast, the positioning of affective emojis does not fit in this general observation. They are regularly sentence-final even in German, as in (17).

(17) a. *Ich wünsche euch einen guten Morgen!* 🥰🤗☺️ [German]
I wish you.PL a good morning SFE
'I wish you all a good morning! (very friendly and blissful tone)'

b. *Ich würde mehr Geld als in meinem Vollzeitjob machen* 😱
I would more money than in my full time job make SFE
'I'd make more money than in my full time job. (shocked tone)' (Twitter)

Some speakers even view modal particles as "verbal emojis," as reflected in the two online remarks below:

> Modal particles are little words that express connotations such as feelings or moods. Because of this, they are also sometimes referred to as "filler words." Basically, they amount to verbal emojis :D (chatterbug.com[10])

[10] `https://chatterbug.com/grammar/german/modal-particles-modalpartikeln` (last visited on 10/22/2022)

> IMO the most important thing to understand about modal particles is that they change mood, not meaning. They are effectively "verbal emojis." (soupsticle on Reddit[11])

The Reddit user in the second quote above further illustrates their point with the similarity between the modal particle *halt* and the shrug emoji 🤷, as in (18).

(18) a. *Das ist **halt** so.* = That's how it is. 🤷

b. *Dann hat er **halt** eine große Nase.* = So he has a big nose, so what? 🤷

The syntactic heterogeneity of affective modal particles and affective emojis in German is most clearly seen when they co-occur in the same sentence, as in (19).

(19) a. *Nachts ist **ja** eine Menge los, dafür muss er **ja*** [German]
at night is MP a lot going on therefore must he MP

tagsüber sehr viel schlafen 😅
during the day very much sleep SFE
'There's a lot going on at night, so he (the speaker's cat) has to sleep a lot during the day. (humorously as-we-all-know tone)'

b. *Wieso ist dir das **denn** so wichtig?* 💅
why is to you that MP so important SFE
'Why is that so important to you? (nonchalantly obliging tone)' (Twitter)

As we can see, the affects in the modal particles and the SFEs add up, just like the situation in Chinese sentences with both SFPs and SFEs (see (5)). In (19a), the modal particle *ja* (which occurs twice) conveys an agreement-seeking, as-we-all-know tone, and the SFE 😅 further adds some minor awkwardness and embarrassment to it (because the speaker's cat sleeps all day long), thus making the overall tone of the tweet humorously fake-serious. Likewise, in (20b) the modal particle *denn* serves to make the question more obliging (and less blunt), while the SFE 💅 adds a nonchalant coloring to the interrogation, thus making the overall tone of the tweet humorously aloof.

To further investigate the syntactic position of affective emojis across languages, I examined posts in nine languages on Twitter and Weibo, as summarized in Table 4. The results show that regardless of the variation in language type and basic word order, affective emojis are invariably sentence-final. See Table 5 for a crosslinguistic illustration (except English, Chinese, and German, which we have seen examples of).

The insensitivity of affective emoji positioning to crosslinguistic word order variation is even more evident in cases where the same content is posted in two languages, as in the Basque and Spanish tweets in (20).

(20) a. *Bilera eta ekitaldi nagusiak bueltan dira Euskaldunan* 😊 [Basque]
Los grandes eventos y las reuniones están de vuelta [Spanish]
en Euskalduna 🎉
'Meetings and big events are back in Basque. (happy and cute tone)'

b. *Bizkaiak egunero zaintzen ditu mendetasun-egoeran dauden* [Basque]
adineko milaka pertsona 👵🏻👴🏻💕

[11] `https://www.reddit.com/r/German/comments/qmit3d/comment/hj9t3f1/?utm_source=share&utm_medium=web2x&context=3` (last visited on 10/22/2022)

Table 4: A crosslinguistic survey of affective emoji positioning

Language	Family	Type	Basic word order	Affective emoji position
Mandarin	Sinitic	isolating	SVO	sentence-final
Japanese	Japonic	agglutinative	SOV	sentence-final
Korean	Koreanic	agglutinative	SOV	sentence-final
English	Germanic	analytic	SVO	sentence-final
German	Germanic	fusional	SOV	sentence-final
French	Romance	fusional	SVO	sentence-final
Irish	Celtic	fusional	VSO	sentence-final
Basque	language isolate	agglutinative	SOV	sentence-final
Hungarian	Finno-Ugric	agglutinative	relatively free	sentence-final

Table 5: Illustration of affective emoji positioning across languages

Language	Example
Japanese	*gozenchū no ame wa dokoni ittandesu ka* 🤔 'Where did the rain in the morning go? (pondering tone)'
Korean	*membeo-deul-i 'hat-gyu'-rago bureum* 🥹 'The members calling him "hot-gyu." (excited fangirl tone)'
French	*C'est réducteur au possible ces fêtes* 😑 'These holidays are as simplistic as possible (frustrated tone)'
Irish	*RT agus fág trácht le bheith san áireamh!!* 😍 'RT and leave a comment to be included!! (enthusiastic tone)'
Basque	*Bilera eta ekitaldi nagusiak bueltan dira Euskaldunan* 😊 'Meetings and big events are back in Basque. (happy and cute tone)'
Hungarian	*Sajnos nem tehetek többet* 😭 'Unfortunately I can't do more. (sad tone)'

Bizkaia cuida cada día de miles de personas mayores en situación de dependencia 👩🏻👴🏻💕 [Spanish]
'Every day, Bizkaia cares for thousands of elderly people in a situation of dependency. (senior-citizen-loving tone)' (Twitter)

In sum, since SFPs and SFEs have clear distinctions in their syntactic behavior, we cannot treat them as elements of the same category.

2.5 Sentence-initial emojis

In the foregoing discussion, I have made the generalization that affective emojis are consistently sentence-final across languages. However, there are also sentence-initial emojis that to some extent encode speaker affects. I discuss three such scenarios in this section and show that none of them is a real counterexample, as they are all qualitatively different from the kind of affective emojis we are concerned with.

2.5.1 Responses to earlier messages

The first type of sentence-initial affective emoji involves normal affective emojis. However, a closer examination reveals that these emojis do not really form a discourse unit with the subsequent sentence but are responses to earlier messages instead. See (21) for an illustration.

(21) a. A: How is she 10 years older than him? She looks 10 years younger 😂.
B: 😂😂 From which angle does she look younger than him? (YouTube)

b. bts.bighitofficial: Left and Right (feat. Jung Kook of BTS) Release
— 😍😍😍 another number one another national anthem 💜✨ (Instagram)

In (21a), B's use of the face-with-tears-of-joy emoji (twice) is an immediate response to A's comment, which B finds hilarious. This usage of affective emojis is reminiscent of interjections, so the double face-with-tears-of-joy emoji can be replaced by words like "hahaha" and "LMAO," and the response would still be felicitous if we remove the subsequent question ("From which angle. . . "). Similarly, (21b) is an Instagram post on the Korean boy band BTS's account together with a fan's comment. The comment begins with an enthusiastic emoji response (three smiling faces with heart-eyes in a row) to the original post. Interestingly, the fan's further comment following the initial response is itself accompanied by a compound sentence-final emoji, which conveys a BTS-loving tone (the purple heart emoji is reserved for BTS in Korean pop culture). The scope difference between the sentence-initial and sentence-final emojis in (21) is intuitively clear and expected. In both cases, some verbal content is added to the discourse first, and some affective content next, with the latter being a response to or a modification of the former.

2.5.2 Creative bullet list icons

Some sentence-initial emojis are bullet list icons. They may be merely for creative visual purposes, as in (22a), or furthermore encode certain speaker attitudes, as in (22b).

(22) a. Ronaldo at the 2002 World Cup:
👕 7 appearances
👌 33.0 touches p/g
⚽ 8 goals
🎯 34 shots/19 on target
✅ 23.5%
⏱️ 69.0 minutes per goal
🔑 13 key passes
📈 7.90 average Sofascore rating

b. ☀️ someday when i comeback to korea, i shall upload many soundclouds
☀️ but is there a spare time between comeback preparation and concert
☀️ ah
☀️ comeback cancel (Twitter)

In (22a), miscellaneous emojis are used to introduce bullet points as well as highlight their themes. These emojis are nonaffective. In (22b), the sun emoji is used by a fan to list some text messages from their idol, and this time the fancy bullet list icon is not only creative but also affective, conveying a warm and affectionate tone.

There are also bullet list icon emojis that are neither theme-specifying nor affective but deictic in nature, in that they directly point to the items they introduce, either literally or figuratively. See (23) for some examples.

(23) a. 👉 Special Shows
👉 Food and blood donations
👉 Upcoming movie posters/videos
👉 Banners,bike rallies and other celebrations
A day with many surprises and celebrations 💥🔥❤️...

b. 📢 The conflict has halted aid deliveries to Tigray...
🗣️ #StopWarOnTigray
🗣️ #EritreaOutOfTigray (Twitter)

Both the point-right emojis in (23a) and the loudspeaker/speaking-head emojis in (23b) are used deictically, drawing readers' attention to the messages they introduce. Also note that the user in (23a) switches to the sentence-final position again when they intend to wrap a text unit in a certain tone—with an ad hoc emoji sequence plus an affective punctuation mark (the ellipsis).

Overall, bullet list icon emojis, whether affective or not, are qualitatively different from the text-accompanying affective emojis we are concerned with, the *major* function of which is tone-setting. As an aside, bullet list icon emojis, being consistently sentence-initial, are not subject to the kind of crosslinguistic word order variation we have observed in affective modal particles either. Their categorial status is beyond the scope of this article but should be part of a general study of emojis.

2.5.3 Decorative frames

Sometimes emojis are used for purely decorative purposes, where they provide a fancy frame for messages or posts and thereby highlight them. See (24) for an illustration.

(24) a. ✨🌿🌸Peaceful Morning🌸🌿✨
🕊️❇️💕Nature Peace💕❇️🕊️

b. 🌱💚DAY 1💚🌱

c. ✨manifesting these two for tomorrow's final✨

d. 🚨EMERGENCY🚨
‼️SWIFTIES LETS VOTE WHILE WAITING FOR SPOTIFY NUMBERS, WE ARE LOSING BADLY‼️ (Twitter)

Being part of a frame, the sentence-initial emojis in (24) are not really sentence-initial but more exactly sentence-surrounding—and they certainly are not only applicable to sentence-level text units either but can enclose any content that the speaker intends to highlight. Thus, they are like fancier versions of the more conventional emphasis asterisks often seen in e-mails, as in (25a). The two types of emphasis markers can also be used together, as in (25b).

(25) a. I know *nothing* about my Indigenous roots.

b. 🚨 **GOOD NEWS ALERT** 🚨 (Twitter)

Compared to the emoji emphasis markers in (24), the asterisks in (25) are less expressive, but the two types of punctuation elements essentially work in the same way. Just like the bullet list icon emojis in Section 2.5.2, these frame emojis are not subject to crosslinguistic variation in positioning either. Beyond their primary function of emphasis scope demarcation, they may additionally encode speaker attitudes. Thus, the emoji frames in (24) respectively convey a peaceful tone, a nature-loving tone, a good-vibe tone, and an attention-craving tone. However, tone-setting is merely a *secondary* effect of frame emojis, again like bullet list icon emojis but unlike the text-accompanying affective emojis we are concerned with in this article, whose *primary* purpose is to set the tone for the text unit they accompany. To avoid ambiguity, we can call the latter *purely affective emojis*.

2.6 Interim summary

Affective emojis in CMC are similar in functionality to affective particles in verbal speech, such as final particles in Chinese and modal particles in German/Dutch. Despite their functional similarity, however, we cannot treat them as the same category in an adequate linguistic analysis of CMC data. First, the two types of affective element often co-occur. And when they do so, they must assume a strict order (SFP $\prec$ SFE). Second, they differ in the open/closed nature of their inventory class, with SFPs being a closed class and SFEs, an open class. In linguistic terms, this suggests that SFPs are more like a grammatical category (for function words), whereas SFEs are more like a lexical category (for content words). Third, purely affective emojis are consistently sentence-final in languages of different families and types, while the positioning of affective particles covaries with the general word order variation across languages. This suggests that SFEs and SFPs are subject to different syntactic rules, which in turn is a clear indication of their distinct categorial status. There are also sentence-initial affective emojis, but those we have observed are either responses to earlier message or have other primary functions (e.g., bullet list–creating, emphasis scope–demarcating) and hence constitute separate uses of emojis. My generalization and theorization in this article are only about purely affective emojis, whose main purpose is to give the text they accompany a certain tone.

The above properties of SFEs present a curious case for linguistic theory. On the one hand, SFEs are functionally similar to SFPs and usually accompany entire text units, which means that their place in the syntactic structure of utterances is in the grammatical zone rather than the lexical domain. It is a basic assumption of modern syntactic theory that the grammatical zone of human language builds on top of the lexical zone. On the other hand, however, the open-class nature of SFEs make them more akin to a lexical category. Further evidence for their lexical status is the frequent conventionalization of their affective senses. For instance, the use of 💅 to convey a nonchalant tone is not predictable from the face value of the emoji, nor is the use of 🤏 by Chinese speakers to convey an onlooker's attitude. Such meaning conventionalization is highly similar to that in content words or idioms. For instance, that *dog* means "a type of four-legged animal" and that *let the cat out of the bag* means "to reveal a secret" are not predictable either and must be learned.

The conclusion we can draw from the foregoing discussion is that SFEs are a *semi-functional-semi-lexical* (henceforth *semilexical*) category. Hence, an adequate linguistic analysis of them should be based on a theory of such categories in general. In the next section, I will introduce such a theory.

3 A formal syntactic theory

Formal syntax is a branch of modern linguistics that approaches the grammatical structure of human language in a formally explicit way. Its origin (in the 1950s) was closely related to formal language theory in computer science (see, e.g., Chomsky 1959), though nowadays most formal syntacticians have shifted the focus of their research to empirically grounded linguistic analysis. My analysis in this section is developed within the Minimalist Program (Chomsky 1995 et seq.). I begin with a general introduction of semilexical elements in human language (§3.1) and then go on to introduce the Generalized Root Syntax theory (§3.2), which is a particular theoretical tool within the Minimalist Program. Finally, I demonstrate how this tool can help us explain the behavior of SFEs (§3.3).

3.1 Semilexicality

Semilexical elements are linguistic elements (mostly words, but also affixes) with both substantive content and grammatical function. By "substantive content," I mean idiosyncratic descriptive content of various sorts. The most familiar classes of words with such content are the major parts of speech (aka lexical categories): Noun, Verb, and Adjective. For instance, *dog*, *cat*, and *bird* are all nouns and can freely substitute for one another in sentences without affecting syntactic well-formedness; they are only considered different words by virtue of their different idiosyncratic content (i.e., they name three different animals). By contrast, purely functional elements, such as the definite article *the* and the infinitive marker *to*, have no such substantive content; they only serve grammatical purposes. Words of the major parts of speech are quintessentially idiosyncratic in meaning, but idiosyncrasy can exist in other word classes too. In other words, purely lexical/functional elements are just two extremes of a continuum, as in Figure 1.

Near the lexical end of the continuum are largely lexical elements that simultaneously perform some grammatical function, such as English light verbs (26). On the other hand, near the functional end of the continuum are largely functional elements that

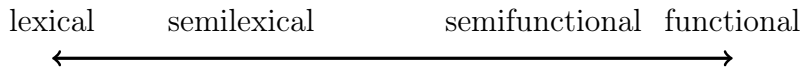


Figure 1: Continuum of lexicality in linguistic elements

simultaneously show some lexical idiosyncrasy, such as Mandarin Chinese conjunctions (27), the usage of which is conditioned by pragmatic factors like the formality of the context.

(26) a. **take** a break, **make** a deal, **do** exercises [English]

b. *hé* 'and (neutral)', *gēn* 'and (colloquial)' [Mandarin Chinese]
yǔ 'and (formal/literary)', *jì* 'and (solemn)'

All three boldfaced words in (26a), termed "light verbs" in linguistics, serve to make verbal predicates out of nouns. It is a special feature of English that such verb-noun collocations select different light verbs, which must be memorized by learners. By comparison, Japanese uses a single light verb *suru* 'do' for all such expressions, as in *kyūkei-suru* 'take a break', *torihiki-suru* 'make a deal', and *undō-suru* 'do exercises' (similarly in Korean, where the general-purpose light verb is *hada* 'do'). In (26b), we can see that instead of a single "and," Mandarin speakers can choose from a number of synonymous conjunctions depending on the context. Thus, the "and" in the book title *Harry Potter and the Philosopher's Stone* is *yǔ*, while that in the ceremony name "Parade commemorating 70th anniversary of the victories of Anti-Japanese War of the Chinese people and the World Anti-Fascist War" is *jì*.

There are also linguistic elements with a more or less even mixture of lexicality and functionality, such as numeral classifiers, which exist in a range of languages. The examples in (27) are from Mandarin Chinese and Japanese.

(27) a. *liǎng-**zhī₁** bǐ* 'two-CL pen' [Mandarin Chinese]
*yì-**zhī₂** cāngshǔ* 'one-CL hamster'[12]
*sān-**zhāng** zhàopiàn* 'three-CL photograph'

b. *ni-**hon** no pen* 'two-CL GEN pen' [Japanese]
*it-**piki** no hamusutā* 'one-CL GEN hamster'
*san-**mai** no shashin* 'three-CL GEN photograph'

The classifiers *zhī₁/hon* are used for long, thin objects; *zhī₂/hiki* (the latter becomes *piki* due to a phonological process) are used for small animals; and *zhāng/mai* are used for thin, flat objects. In classifier languages like Chinese and Japanese, different nouns require different classifiers, but all classifiers share the same grammatical function—they all turn mass concepts into countable units. Classifiers lie somewhere near the midpoint of the continuum in Figure 1, for they are more functional than semilexical elements (with their fundamental status in the grammar being functional) and more lexical than semifunctional elements (with their idiosyncratic content being more substantive than just pragmatic conditioning).

For the purpose of this article, I will simply use "semilexicality" as a cover term, without further distinguishing the fine-grained subtypes above. The phenomenon as

[12]The two classifiers *zhī₁* and *zhī₂* are etymologically unrelated and also written differently in the Chinese script, respectively as 支 and 隻.

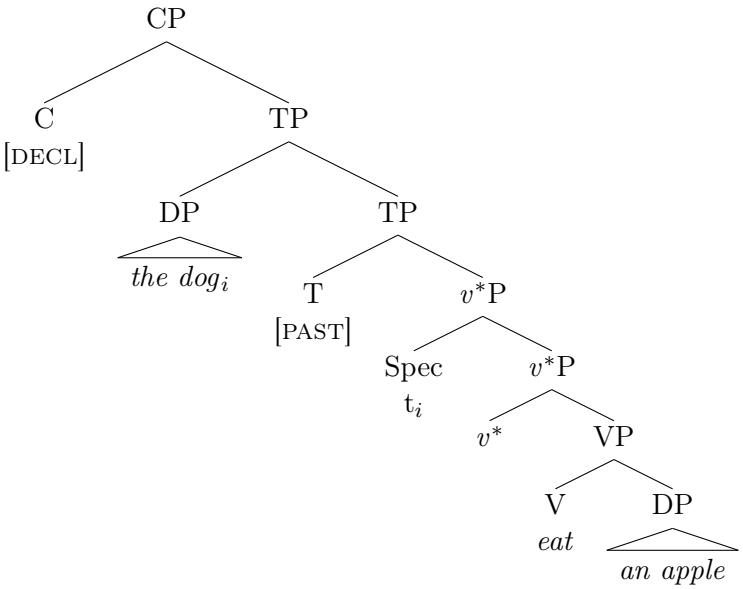
a whole is receiving increasing attention in theoretical linguistics (see Song 2021 for a typological discussion).

3.2 Generalized Root Syntax

The semilexicality phenomenon is a challenge for formal syntax, where syntactic categories are either lexical or functional, with no third possibility. This has to do with the way the lexicon and the syntax are theoretically connected. Simply put, syntactic derivation in the Minimalist Program starts with a lexical base, to which multiple layers of functional extension are added. Take the simple sentence in (28a) for example. Its syntactic structure is diagrammatically represented in (28b) (with some simplification for expository convenience). The tree diagram can be read from the top down as follows: "The sentence in (28a) is a CP consisting a functional head C and a TP complement selected by that head; TP consists of ...."

(28) a. The dog ate an apple.

b.



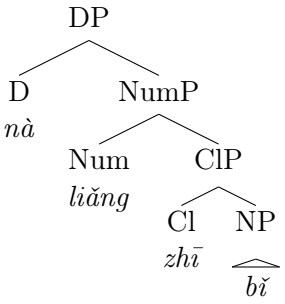
The syntactic tree in (28b) shows the formal derivation of a clause, which consists of a single lexical category V plus three functional categories v^*, T, and C. These respectively serve to specify the agentive subject (i.e., the doer),[13] the tense (past), and the type (declarative) of the clause. The lexical category V itself, on the other hand, is only responsible for introducing the core predicate (an eating activity) and its direct object (*an apple*). Leaving many technical details aside (e.g., the Spec node and the two DP triangles), we should notice that a syntactic category or "head" in the tree is either lexical or functional. There is simply no other possibility.

Now, let's turn to a phrase with a typical semilexical element, as in (29).

(29) a. *nà liǎng zhī bǐ* [Mandarin Chinese]
those two CL pen
'those two pens'

[13]The subject subsequently moves to a higher position by transformation, which is conveniently indicated by a t (for "trace") and a coreferential index i in (28b).

b.

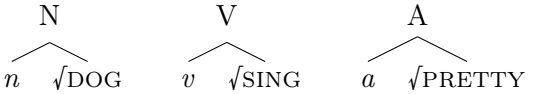


The Mandarin Chinese phrase in (29a) has one lexical category N, which forms its base, plus three functional categories: Cl (for the classifier), Num (for the numeral), and D (for the determiner). Crucially, the Cl head is *functional* in the syntactic system despite the semilexical nature of actual classifiers. It is impossible to reflect the semilexicality at the categorial level of the formal representation.

Linguists have noticed the above theoretical problem and also made attempts to bypass it. A representative solution, which has been independently proposed several times in recent years, is to resort to a *root-based analysis* (see, e.g., Acedo-Matellán and Real-Puigdollers 2019, Song 2019, and Pots 2020). Root is a notion from an influential offshoot of generative syntax known as Distributed Morphology (henceforth DM; Halle and Marantz 1993 et seq.; see Harley and Noyer 1999 for a concise introduction), which treats word structure as syntactic structure and assumes a single generative engine for human language (i.e., the syntax). In DM, the root formalizes the idea of a categoryless (aka acategorial), purely lexical element, which does not even have a major part of speech. A key hypothesis of DM is that the atoms of syntactic derivation are acategorial roots (aka l-morphemes) and purely functional categories (aka f-morphemes) instead of ready-made words. On this hypothesis, what used to be considered minimal syntactic objects, most typically bare words of the major parts of speech, are given a further layer of subatomic analysis, as in (30).

(30) a. dog, sing, pretty

b.

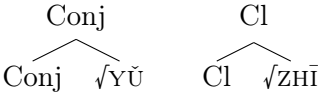


(n = nominalizer, v = verbalizer, a = adjectivizer)

The three roots in (30), which are typeset in small capital letters and preceded by a square root symbol, are void of categorial information. They only get "categorized" by being merged with a special functional head, called a *categorizer*. Thus, the n-√DOG merger yields a noun *dog* based on the root √DOG. If we merge the same root with a different categorizer, we may get a different word of a different category. For instance, the v-√DOG merger yields a verb meaning "to follow very closely" or "to ask constantly." Of course, which categorizer-root merger yields what word—or whether it corresponds to an existing word at all—is a matter of language-specific lexicalization. Thus, while the root √DOG is the base of both a noun and a verb (and apparently also an adjective, as in *dog French*), the root √BOY is only the base of a noun in current English—though a verb or an adjective *boy* is a theoretically possible word and may well be coined. The DM categorization schema just formally represents the intuition that each content word (in a given context) has a specific syntactic category plus some idiosyncratic substantive information.

In standard DM, the root categorization tool is only reserved for content words. However, as Acedo-Matellán and Real-Puigdollers (2019), Song (2019), and Pots (2020) among others argue, it may be applied to semilexical words too. The logic is simple: when the categorizer is not a major-part-of-speech f-morpheme but an ordinary functional category, its merger with a root essentially yields a function word with some idiosyncratic content (contingent on language-specific lexicalization). Song (2019) explicitly distinguishes this extended use of the DM tool from its original use by calling the former Generalized Root Syntax.[14] See (31) for an illustration.

(31) a. *yǔ* 'and (formal/literary)' [Mandarin Chinese]
zhī 'classifier for long, thin objects'

b. 

As we can see, Conj and Cl are both normal functional heads, but when they are respectively supported by the roots √YǓ and √ZHĪ, we get a conjunction and a classifer with additional idiosyncratic content (which, in the latter case, is just the usually understood idiosyncrasy of the classifier). These roots can in theory merge with other functional categories to yield other words, and this is indeed the case. Thus, √YǓ can also be categorized into a preposition meaning "with" (32a), and √ZHĪ can also be categorized into a verb meaning "prop up, put up" (32b).

(32) a. *tāmen xīwàng* ***yǔ*** *jiārén yìqǐ guò-jié* [Mandarin Chinese]
they hope with family together spend-holiday
'They hope to spend the holiday with their families.'

b. *máfan nǐ bǎ sǎn* ***zhī****-kāi yíxià*
bother you DISP umbrella put.up a.bit
'Could you please put up the umbrella for me?'

3.3 Sentence-final emojis, formally

The same analytical method can be applied to SFEs. That is, we can separate their shared function (i.e., marking speaker affects) from their specific content (i.e., the affects) by encoding the former in a functional head, call it E (for "emotion"), and the latter, in an acategorial root. What is unique about the root of an SFE is that it is visual-digital instead of verbal-linguistic.

Since CMC is not confined by the conventional modalities of communication (e.g., oral-auditory, visual-manual), the theoretical space of roots—and thereby words in a broad sense—can be extremely large. The digital modality makes available a wide variety of elements (e.g., icons, pictures, GIFs) that may be readily recycled for communicative purposes. Since such recycled visual elements each associate a form with a (contextualized) meaning, their role in CMC is just like that of words in face-to-face (or voice-to-voice) communication—though clearly the shape of words is much more versatile in CMC. We can call the pre-recycling visual elements *digital roots* and call the communicative recycling procedure itself *digital categorization*. I illustrate this procedure in (33) with the

[14]Borer (2013) has a similar proposal in a different (non-DM) theoretical setting.

nonchalant-tone SFE 💅 (the example sentence is repeated from (9)). The subscript √ notation in (32b) indicates that the abstract category E is now supported by the idiosyncratic information of a root.

(33) a. had 'hug' been a little more second longer, she would've elbowed one of these queens out. just saying 💅 (Twitter)

b. $E_{\surd}$ ⇐ `nail polish image used affectively (nonchalant tone)`

⇐ `nail polish image`

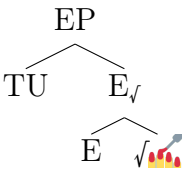
The image 💅 itself does not necessarily denote nonchalance. At face value, it is just a nail polish icon, which may well just denote a nail-polishing activity in a different context, as in (34).

(34) Enroll for various nail courses at Riva and pursue your dreams of becoming a nail 💅 technician. (Twitter)

What triggers the nonchalance reading of 💅 in examples like (33a), therefore, is the affective categorial context—or in formal linguistic terms, the functional category E. And that reading itself is a result of conventionalization, just like the meaning of any content word or idiom. In defense of Root Syntax, Marantz (1995) famously asserted that *cat* was a phrasal idiom. By the same token, we can say that each affective emoji is a tiny idiom in the CMC lexicon, because we cannot predict its affective reading with full confidence (even for simple smileys like 🙂, which is more passive-aggressive than friendly in current usage) but must learn it as we learn any other new word.

Following the categorization step in (33b), the root-supported $E_{\surd}$ can project its own phrase structure like any other functional category can. This gives us the structure in (35), where TU stands for "text unit."

(35)



The root-supported $E_{\surd}$ merges with the text unit it accompanies and labels the product of this merger EP. In other words, E functions like an emotional wrapper around the text unit it accompanies.

The root-based syntactic analysis above makes several immediate predictions about the behavior of affective emojis, which exactly correspond to what we observed in Section 2. First, since the affective meaning triggered by the E-√ merger is a result of language-specific conventionalization, the same emoji form may have different meanings in different languages/cultures or for people of different generations. In other words, emojis are not a universal language, contrary to a popular impression. The above-mentioned simple smiley 🙂 is a good example of cross-generational variation. The shift in its affective meaning is similar to that in the meanings of content words like *awful* 'impressive→extremely bad' and *gay* 'joyous→homosexual'. An example of cross-cultural variation is the aforementioned Weibo emoji 😏, which is popularly used in China to express an onlooker attitude

but does not have this usage in other cultures. Similarly, the dog-head emoji 🐶 (popularly named "doge"), which does not exist in Unicode but does on a number of Chinese platforms (e.g., 😸 on WeChat, 🐺 on Douyin), has more or less become *the* emoji for sarcasm in China, as in (36).

(36) A Weibo user posted that they had brought a lot of food to the quarantine hotel, and someone replied:

zěnme méi bǎ kōngqì zháguō dàishàng 🐶 [Mandarin Chinese]
how not.have DISP air fryer bring.along SFE
'How come you haven't brought along your air fryer? (sarcastic tone)' (Weibo)

With the dog-head emoji, it is clear to Chinese speakers that the question is not genuine but sarcastic (though not really hostile).

The second prediction of the analysis is that affective emojis are pheripheral word-order-wise. They can be either to the left or the right of the text unit they accompany, but cannot be in its middle. Formally speaking, this is because the position of $E_{\surd}$ is outside of the TU position in (35). The conversion of hierarchical syntactic structures to linear strings is rule-based, and there are only two linearization possibilities for the tree in (35): TU $\prec$ $E_{\surd}$ or $E_{\surd}$ $\prec$ TU. This means that there can be truly sentence-initial affective emojis beyond the marginal cases in Section 2.5, which is a point that needs further attestation. For now, we can probably explain the predominantly sentence-final positioning of affective emojis by the content-before-emotion communicative habit of Internet users and the left-to-right directionality of the scripts in our data. This means that in languages with right-to-left scripts, affective emojis will show up to the left of the text they accompany. This is indeed the case, as evidenced by the Hebrew example in (37).

(37) 💙 אוהב אתכם [Hebrew]
💙 ohev etkhem
love you.PL SFE
'Love you. (affectionate tone)' (Twitter)

The blue heart emoji in (37), despite its geometric initiality, is logically sentence-final. Interestingly, the translation functionality of Twitter would automatically switch the geometric positioning of emojis too when translating from Hebrew to English.

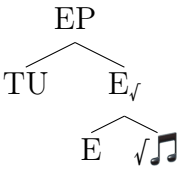
The third prediction of the root-based analysis is also about linearization. The above-mentioned two possibilities to order TU and $E_{\surd}$ are still based on requirements of the oral-auditory modality—in particular, the requirement that linguistic structures must unfold linearly in time. However, such strict linearity is not a requirement of CMC, since the channel of externalization (i.e., the digital screen) is two-dimensional. Thus, the positioning of $E_{\surd}$ with respect to TU ought to have more flexibility than what we have seen so far. In theory, the EP structure in (35) can be externalized in any way that does not interpolate $E_{\surd}$ *inside* TU. Thus, we can view $E_{\surd}$ and TU as being placed in two different layers (as in Photoshop), which may be organized in whatever way the 2D screen allows for: horizontally, vertically, or with overlay. This extended view of EP linearization makes it possible to give affective emojis and memes a unified formal analysis. See (38) for an illustration.

(38) a. b. c.

The three memes in (38) are respectively in English, Chinese, and Spanish, and they each externalize EP in a different way: vertically in (38a), with TU-over-$E_{\surd}$ overlay in (38b), and with $E_{\surd}$-over-TU overlay in (38c). Note that (38c) does not really involve interpolation despite its separation of the content of TU on two sides of the affective image, because when reading the meme, we still read the text as *Ojalá TODO vaya bien* instead of *Ojalá TODO* 🫣 *vaya bien*. Besides, the distributed positioning of the Spanish sentence is not just with respect to the image either, but is more exactly with respect to the entire canvas (to further use the analogy with Photoshop) and everything contained in it, as is evidenced by the larger-than-usual space between *Ojalá* and *TODO*. In other words, what we see in (38c) is a case of *geometric* rather than *logical* interpolation. The root-based analysis merely predicts the impossibility of the latter but not that of the former, for geometric positioning has more to do with graphic design than with linguistic externalization.

Last but not least, the above analytical framework allows us to further expand the scope of affective elements. The digital modality is more flexible than naturally evolved biological modalities not only in terms of image type (icons, emojis, GIFs, etc.) and linearization possibility, but also in terms of the more general "filetype" of the affective element. So far, we have limited our discussion to affectively recycled visual elements, but on the Internet, audio elements may be recycled too. This is what happens in Instagram posts or "stories" with background music. The linguistic structure of such multimedia posts is exactly the same as that of affective emojis/memes, as in (39), where I use 🎵 to denote some audio element.[15]

(39)



In sum, the digital modality provides a much bigger stage for the affective modification of linguistic expressions than biological modalities, of which affective emojis are just a particular manifestation. The root-based analysis presented in this section is suitable for the affective recycling of all kinds of multimedia material.

[15]More often than not, Instagram posts with background music also have background images. On the current analysis, this requires the root part of the structure to be a multimedia compound, which is similar to the situation in the emoji sequences we have seen on p.8.

4 CMC linguistics

In Section 1, I asked two general questions: one about the cognitive nature of CMC, and the other about tools from modern linguistics that are applicable to it. My investigation of affective emojis in Sections 2–3 reveals that there is indeed some substantial cognitive difference between oral languages and CMC. The difference mainly lies in the nonbiological nature of the digital modality, whose flexibility and extensibility are far beyond the capacity of naturally evolved modalities of communication. We have discussed visual and audio affective elements in this article, but as newer technologies arise, there will certainly come newer types of communicative elements too, such as elements of virtual reality or the metaverse.

The unique features of CMC requires us to rethink the relation between language and writing/typing in the 21st century. CMC is clearly still built on conventional linguistic content, either written/typed or spoken/recorded. But the ever-increasing information processing and transmitting power of the computer enables users to further modify the linguistic content in unprecedented ways. It is such computer-mediated modification that requires linguists' careful investigation. The reason is that such modification counts as an "interface" issue of the digital modality.

In generative linguistics, especially in the Minimalist Program, interface legibility conditions are taken to be a major driving force and gauge of success for linguistic theory. These are conditions that a generative theory of human language must meet to ensure that the structures it generates are legible in the cognitive systems that the language faculty interfaces with. For instance, to make sure that linguistic structures are legible by the sensorimotor system, some algorithm must apply to convert them into linear strings. An influential proposal in this regard is Kayne's (1994) Linear Correspondence Axiom. On the other hand, to make sure that the linguistic structures are legible by the conceptual-intentional system, some operations must apply to remove uninterpretable features from them, such as features of grammatical case (e.g., accusative) and agreement (e.g., first-person singular). Quite a few key operations of the Minimalist Program (e.g., Agree, Delete) are motivated by this legibility condition.

Given the fundamental significance of interface conditions, linguists must ask themselves whether the same conditions still apply in the case of CMC. This is a legitimate question, since each interface presumably has its own legibility conditions. My case study in this article demonstrates that the syntax-pragmatics interface is strongly influenced by the change of modality, because CMC makes available a myriad of communicative elements (e.g., affective emojis) that take effect at the pragmatic level. Beyond the immediate scope of this article, however, I think the big-picture question we need to ask is:

- How must linguistic theory adapt itself to the cyber-digital interface?

By the cyber-digital (henceforth C-D) interface, I mean the interface between the language faculty and the computer-and-network system that CMC relies on. Note that this interface is an unusual one from a linguistic perspective, because while all other linguistic interfaces are within the confines of the mind, the C-D interface is not—unless the computer is viewed as an extension of the human mind. The unusualness of the C-D interface means that to answer the question above, we must first answer the question below:

- How likely is it for the cyber-digital system to replace the sensorimotor system as an alternative modality of language externalization in the human world?

As things currently stand (in the early 21st century), the likelihood is quite small. But if there comes a day when the answer to the last question becomes a positive Yes (i.e., when cyborgs no longer only exist in fiction), then CMC linguistics should definitely become an official branch of linguistics, even if grapholinguistics still remains marginal.

As far as I am concerned, until the legibility conditions of the C-D interface are ascertained, perhaps the safest theoretical linguistic tools to use in the study of CMC—or more exactly the CMC-specific part of CMC data (e.g., emojis)—are none other than the most basic ones—those that are *not* designed to meet the generativists' interface conditions but are independently needed by any adequate theory of human language. In particular, I can think of the following three tools, the first two of which I have already used in my case study of emojis:

1. The basic combinatorial operation that builds complex linguistic units out of simpler ones: This operation lives under various names in different theoretical frameworks. It is called "Merge" in the Minimalist Program, which is formally just set formation: $\text{Merge}(A, B) = \{A, B\}$.

2. The recycling of existing materials for new purposes: This is essentially what Generalized Root Syntax is about, where miscellaneous root materials may be recycled to support and enrich abstract functional categories. Depending on the nature of the particular functional category, this may correspond to "grammaticalization" or "lexicalization" in traditional linguistic terminology.

3. The compositional interpretation of syntactic structures: This is what another major branch of theoretical linguistics, formal semantics, is about. Since the formal tools in compositional semantics (e.g., the lambda calculus, first-order logic) are not limited to the analysis of natural languages but are generally applicable to any symbolic system, they can certainly be used to represent the semantics of CMC data too.[16]

Thus, the safest tools to use in CMC linguistics, for the time being, are either tools that are not motivated by interface conditions or tools that are not designed for the analysis of natural language alone. On that note, the first two tools above are perhaps not entirely natural language–specific either but may be viewed as the manifestation of some domain-general strategies in the language domain: Merge qua set formation is obviously needed in many cognitive domains (e.g., mathematics), while the recycling of existing materials for new purposes is essentially just assigning old materials new categories, and categorization is one of the most fundamental coginitive processes underlying human intelligence, which clearly is domain-general too. On the other hand, many familiar generative linguistic tools (e.g., movement, phase-based spell-out) are not entirely safe due to their oral language–specific nature, or more generally due to their strong association with the legibility conditions of the sensorimotor interface. I have refrained from using such tools in my analysis of affective emojis.[17]

[16]See Song (2022) for a compositional semantics for the emoji syntax proposed in this article.

[17]I am not denying the utility of domain-specific tools in the study of CMC data in general but merely cautioning against their application in the study of the CMC-specific part thereof, such as emojis. One can certainly use operations like movement in the analysis of the linguistic expression basis of CMC data; i.e., the TU part of (35). A caveat here is that the separation of CMC data into a CMC-specific and a non-CMC-specific part might entail a more complicated (and potentially multiparty) interface relation between the various systems involved in CMC linguistics.

The "safe" nature of domain-general tools is reminiscent of what Chomsky (2005) has designated the "third factor" in language design—namely, principles that are not specific to the language faculty, such as principles of data analysis or processing and principles of structural architecture and efficient computation. According to Chomsky, such principles are not motivated by the need of the language faculty alone but are nevertheless an indispensable part of the growth of language in the individual. It seems therefore that the part of the generative linguistic tool kit suitable for research on CMC (again until its interface conditions become clear) is just the set of tools that can be cast as third-factor strategies.

5 Conclusion

In this article, I presented a formal linguistic study of affective emojis (aka sentence-final emojis) in CMC data and laid out some preliminary thoughts on CMC linguistics. The point of departure for my case study is the syntactic analysis of such emojis in Song (2019). While I have inherited and revised Song's (2019) root-based analysis, I have objected to his unified treatment of sentence-final particles in oral languages and sentence-final emojis in CMC based on three arguments (§2). My revised analysis (§3) separates CMC data with affective emojis into a non-CMC-specific part (i.e., the linguistic text) and a CMC-specific part (i.e., the emoji), with the latter functionally wrapping around the former and thereby setting a tone for it. A merit of this analysis is that it can be directly applied to other CMC-specific affective elements too, such as memes and background music. More generally, the analysis is suitable for any affective modification of linguistic expressions in the digital modality of communication. The special nature of the digital modality has nontrivial ramifications for CMC linguistics (§4). Until the legibility conditions of the cyber-digital system are ascertained, the safest linguistic tools to use in research on CMC-specific phenomena are the domain-general ones, or the ones that can be cast as Chomsky's (2005) "third factor" strategies. I will explore the legibility conditions of the C-D interface as well as the interface relation in CMC linguistics to future research.

References

Acedo-Matellán, V. and Real-Puigdollers, C. (2019). Roots into functional nodes: Exploring locality and semi-lexicality. *The Linguistic Review*, 36(3):411–436.

Biberauer, T. and Roberts, I., editors (2013). *Challenges to Linearization*. De Gruyter Mouton, Berlin.

Borer, H. (2013). *Taking Form*, volume 3 of *Structuring Sense*. Oxford University Press, Oxford.

Carey, J. (1980). Paralanguage in computer mediated communication. In *Proceedings of the 18th Annual Meeting of the Association for Computational Linguistics*, pages 67–69.

Chao, Y.-R. (1968). *A Grammar of Spoken Chinese*. University of California Press, Berkeley.

Cheng, S.-P. and Tang, S.-W. (2022). Syntax of sentence-final particles in chinese. In Huang, C.-R., Lin, Y.-H., Chen, I.-H., and Hsu, Y.-Y., editors, *The Cambridge Handbook of Chinese Linguistics*, Cambridge Handbooks in Language and Linguistics, page 578–596. Cambridge University Press, Cambridge.

Chomsky, N. (1957). *Syntactic structure*. Mouton, The Hague.

Chomsky, N. (1959). On certain formal properties of grammars. *Information and Control*, 2(2):137–167.

Chomsky, N. (1995). *The Minimalist Program*. MIT Press, Cambridge MA.

Chomsky, N. (2001). Derivation by phase. In Kenstowicz, M., editor, *Ken Hale: A Life in Language*, pages 1–52. Oxford University Press, Oxford.

Chomsky, N. (2005). Three factors in language design. *Linguistic Inquiry*, 36(1):1–22.

Durrell, M. (2021). *Hammer's German Grammar and Usage*. Routledge Reference Grammars. Routledge, Taylor et Francis Group, 7 edition.

Fehringer, C. and Cornips, L. (2019). The use of modal particles in Netherlandic and Belgian Dutch imperatives. *Journal of Germanic Linguistics*, 31(4):323–362.

Gawne, L. and McCulloch, G. (2019). Emoji as digital gestures. *Language Internet*, 17(2). `https://www.languageatinternet.org/articles/2019/gawne/index_html`.

Grosz, P. G., Greenberg, G., De Leon, C., and Kaiser, E. (2021). A semantics of face emoji in discourse. Manuscript, to appear in *Linguistics & Philosophy*. `https://ling.auf.net/lingbuzz/005981`.

Halle, M. and Marantz, A. (1993). Distributed morphology and the pieces of inflection. In Hale, K. and Keyser, S. J., editors, *Essays in Linguistics in Honor of Sylvain Bromberger*, number 20 in The View from Building, pages 111–176. MIT Press, Cambridge MA.

Haralambous, Y. (2020). Grapholinguistics, TEX, and a June 2020 conference. *TUGboat*, 41(1):12–19.

Harley, H. and Noyer, R. (1999). Distributed morphology. *Glot International*, 4(4):3–9.

Kayne, R. (1994). *The Antisymmetry of Syntax*. MIT Press, Cambridge MA.

Li, C. and Thompson, S. (1981). *Mandarin Chinese: A Functional Reference Grammar*. University of California Press, Berkeley.

Maier, E. (2021). Emojis as pictures. Manuscript, to appear in *Ergo*. `https://ling.auf.net/lingbuzz/006025`.

Marantz, A. (1995). *Cat* as a phrasal idiom: Consequences of late insertion in distributed morphology. Manuscript, Massachusetts Institute of Technology.

Morita, E. (2018). Sentence-final particles. In Hasegawa, Y., editor, *The Cambridge Handbook of Japanese Linguistics*, Cambridge Handbooks in Language and Linguistics, pages 587–607. Cambridge University Press, Cambridge.

Paul, W. (2014). Why particles are not particular: Sentence-final particles in Chinese as heads of a split CP. *Studia Linguistica*, 68(1):77–115.

Pierini, F. (2021). Emojis and gestures: A new typology. In *Proceedings of Sinn und Bedeutung*, volume 25, pages 720–732.

Pots, C. (2020). *Roots in Progress: Semi-lexicality in the Dutch and Afrikaans verbal domain*. PhD thesis, KU Leuven.

Potts, C. (2005). *The logic of conventional implicatures*. Oxford University Press, Oxford.

Saussure, F. d. (2011). *Course in General Linguistics*. Columbia University Press, illustrated, revised edition. Ed. by P. Meisel and H. Saussy. Translated by Wade Baskin.

Song, C. (2019). *On the formal flexibility of syntactic categories*. PhD thesis, University of Cambridge.

Song, C. (2021). A typology of semilexicality and the locus of grammatical variation. Talk at the 9th International Conference on Formal Linguistics (ICFL9), Nov 5–7, Fudan University (online). `https://www.juliosong.com/doc/Song2021ICFL9.pdf`.

Song, C. (2022). Sentence-final particle vs. sentence-final emoji: The syntax-pragmatics interface in the era of CMC (extended version). SyntaxLab talk, Jun 28, University of Cambridge (online). `https://www.juliosong.com/doc/Song2022SynLabJun.pdf`.

Sun, X. (1999). *Interjections in Modern Chinese*. Language and Culture Press, Beijing.

Index