

Going ✌️ Lexicon? The Linguistic Status of Pro-Speech Emojis

John David Storment
Stony Brook University

Abstract: Emojis, introduced in the US in 2011 and now ubiquitous, are a set of iconic expressive symbols that are incredibly widespread in computer-mediated communication (CMC), especially among young people. The majority of linguistic research on emojis focuses only on the semantics; however, emerging data suggest that emojis are far more linguistically interesting than merely their semantic contributions to a sentence. Data from Twitter demonstrate that emojis can actually appear as contentful morphological units that behave according to regularly predictable morphosyntactic rules. In this paper I analyze data from several languages including English, German, and Spanish and reach several conclusions about emojis that appear as lexical items in sentences. First, these emojis are not merely replacements in text for existing words in a language, but rather they represent morphological stems to which inflectional and derivational affixes can be productively added. Second, emojis can undergo morphological changes such lexicalization and grammaticalization. Within pro-speech emojis, I also differentiate two sub-categories of emoji according to how they are interpreted: first, emojis which seem to be a code for a different lexical item, and second, emojis whose interpretation is entirely iconic, for which it is beneficial to say some of them are distinctly represented in the language user's lexicon.

Keywords: emojis, CMC, morphology, affixation, stems, inflection, derivation, lexicalization, grammaticalization, iconicity, pro-speech gestures

1. Introduction

Emojis (often bare plural *emoji*) are a relatively new form of visual communication that exist as an option on most digital keyboard layouts. They are symbols that depict objects, signs, ideas, and facial expressions or smileys. People from all backgrounds and in various languages use emojis in daily communication online and in text messages, and to varying degrees. Use of emojis is now so frequent that they can be seen in advertisements (Knudson 2017), in print (Evans 2017), on clothing, and even personified in film (Leondis 2017).

Emojis have been researched by computer scientists, psychologists, sociologists, and linguists (Bai et. al. 2019). The majority of linguistic research on emojis has focused on the semantics – how the meaning an emoji has can compose into the truth conditions of a sentence. This research follows the recent trend of asking the questions *how can extralinguistic information contribute to the information conveyed in language?* and *can nonlinguistic meaning be modeled with tools from formal semantics?* This research contributes to the growing body of literature on so-called super semantics (Schlenker 2019). Yet, emojis bring up a multitude of interesting linguistic questions. The sociolinguistic use of emojis has been documented (Moschini 2015), and some have asked whether or not emojis have a syntax (Cohn, Engelen, & Schilperoord 2019).

The focus of this paper, however, is the morphology of emojis that appear inside of sentences, exemplified by the following well-known sentence.

(1) I ❤️ NY

This example may seem straightforward and relatively uninteresting, but there is a wealth of data from Twitter and other websites that show very interesting morphosyntactic properties of emojis that appear in syntactic positions, which is confirmed by the judgments of proficient emoji users. From these data and their analysis I show that these emojis are predictable and cohesive morphological units that can be predicted with regularity, and that can participate in morphosyntactic processes such as inflection, derivation, lexicalization, and grammaticalization. I even raise the question of whether some of these emoji words have entered some language users' lexicons as independent units.

The paper is structured as follows: in Section 2 I outline a brief history of emojis and their use. In Section 3 I describe three functions that emojis can have in a written utterance. In Section 4 I expand in detail upon the facts surrounding the distribution of emojis that appear within sentences instead of words in three languages: English (4.1.1.), German (4.1.2.), and Spanish (4.1.3.). I discuss data of these emojis taking inflectional (4.1.) and derivational (4.2.) affixes. I then discuss morphological regularization (4.3.) and summarize the findings (4.4.). In Section 5 I show that these emojis are sensitive to processes of morphosyntactic change such as lexicalization (5.1.) and grammaticalization (5.2.). In Section 6 I discuss a possible two-way classification of these emojis based on their interpretations. Section 7 concludes.

2. Some Background on Emojis

The first emoji-like elements (in-text ideographic symbols not composed of preexisting characters) emerged in Japan in the late 1990s to convey the meanings of facial expressions seen in manga (Danesi 2017). These symbols eventually evolved into what we know as emojis. The

word *emoji* comes from Japanese 絵 *e* ‘picture’ and 文字 *moji* ‘character’, and was first introduced by emoji inventor Shigetaka Kurita in 1999. The resemblance this word bears to English *emoticon* is coincidence, as *emoticon* is a portmanteau of *emotion* and *icon* that is attested since the mid-1980s. The emoji keyboard was introduced for Apple digital keyboards on the iPhone in 2008 with iOS 2.2 and is now a standard feature on almost all smartphones, tablets, and computers (Evans 2017).

Using ideograms in writing to convey emotion is no new feature of written language, though. The first attested smiling face – or smiley – in a text is found in a financial record in the town of Trenčín, Slovakia next to the author’s signature (Ladislavides, 1635). The symbol was used identically to one of the ways in which contemporary emojis are used, to indicate the author’s positive feelings about the preceding text (Votruba 2018).

The first recorded use of the emoticons :-) and :-(in CMC, a predecessor to emojis that are still frequently used, occurred in 1982 in a computer forum (Fahlman 1982). Though some have proposed that the expressive symbol consisting of a colon followed by a closed parenthesis is attested since the seventeenth century (Stahl 2014), these instances are more than likely coincidental.

The original Apple emoji keyboard contained 471 emojis, but there are 4,084 at the time of writing this paper. Each emoji has its own Unicode entry, keywords, and a short name (“grinning face”, “squid”, “flag of Australia”, etc.), which I will be subsequently referring to as emoji’s name or label. The original was comprised of 32 smileys, a number of pictograms and ideograms depicting various animals, objects, and abstract symbols, as well as the flags of ten countries. The current emoji keyboard contains over 100 smileys, over 100 pictograms of humans and human body parts in six different skin colors, hundreds upon hundreds of miscellaneous pictograms and ideograms, flags for every country in the world, and flags for the gay and trans communities (Emojipedia). New emojis are added continuously with each Apple software update. Which emojis become popular and the ways in which emojis are used is somewhat unpredictable as there is considerable variation in the frequency of emoji use and the ways in which they are used.

3. Functions of Emojis in Sentences

Emojis have multiple possible configurations of where they can go in a sentence, and these positions correspond to different linguistic functions. In this section I identify three broad categories of where emojis can go. I borrow terminology from the domain of super linguistics and the body of research on the semantics of gestures (Schlenker 2018) to name the categories *post-speech emojis*, emojis that follow an utterance (3.1.), *co-speech emojis*, emojis that immediately follow and are directly associated with one specific word or constituent embedded in a sentence (3.2.), and *pro-speech emojis*, which are emojis that are actually projected in the sentence as words¹, and will be the main focus of the morphological analysis of this paper (3.3.).

¹ ‘Words’ here is an oversimplification of the actual morphosyntactic status of pro-speech emojis, which will be fully fleshed out later on.

3.1. Post-Speech Emojis

The vast majority of the current literature on formal linguistics of emojis is about the semantics of sentences containing post-speech emojis (Grosz, Kaiser & Pierini 2021; Maier 2020; Pasternak & Tieu 2020; Pierini 2021), in which the semantics of these emojis are often analyzed in the same way as the semantics of gestures that follow an utterance. Post-speech gestures are separate discourse units linked to the previous utterance via an anaphoric link and a discourse relation; in other words, they behave like appositive relative clauses (Schlenker 2018). Within post-speech emojis, Grosz, Kaiser, & Pierini (2021) distinguish between two types: face emojis (2-3) and activity emojis (4).

- (2) a. Did you see that guy? 😍 (Grosz et al 2022)
b. That fried chicken sandwich they make 😋
c. If a movie is violent, Alex hates it 😞

These face emojis in (2) are expressive elements that provide information on the attitudes that the speaker feels toward the proposition in the text. For example, the “smiling face with heart-eyes” emoji in (2a) indicates that the speaker feels infatuation about the aforementioned guy. In (2c), the “worried face” emoji expresses the speaker’s negative emotions about Alex’s hatred of violent movies. The examples in (2) are all dependent, which means that the interpretation in the text influences the interpretation of the emoji since the emoji comments on the proposition of the text, and this is the focus of Grosz et al (2022), but there are examples where the emoji is independent as it offers no comment on the speaker’s attitude toward the text. See these below.

- (3) a. How did the interview go? 😊 (Grosz et al 2022)
b. How are you coping? 😞

These examples are the most clearly similar to gestures and facial expressions in terms of how their meaning composes into the sentence. The activity emojis in (4), while also appearing after the sentence, serve a different function. For Grosz, Kaiser, & Pierini (2021) these are event descriptions whose semantics are comparable to free adjuncts.

- (4) a. Arsenal really impressed me ! ⚽ (Grosz, Kaiser, & Pierini 2021)
b. Getting ready for tomorrow! 🏆
c. My job is pretty fun 🪄

Post-speech emojis do not represent syntactic constituents, and it would be erroneous to suggest that they are projected in the syntax of the sentences they modify, just as no one suggests that post-speech gestures are projected syntactically. This is not to say, however, that no emojis can be projected syntactically.

3.2. Co-Speech Emojis

Less studied are cases of what I dub co-speech emojis, where emojis directly follow (often without spacing) a word or phrase that they modify.

- (5) I can build a house 🏠 rebuild a car 🚗 and dig your grave!!!
<https://twitter.com/K7711mcc/status/1601683049695309824>
- (6) Breathe on me while I sleep 😴 tonight Lord!
<https://twitter.com/lcircum/status/1150465737787617280>
- (7) Trans 🇺🇸 people are human!!
<https://twitter.com/TsMadisonat1/status/1377813662187057157>
- (8) drink 🍺 like Peter // hate like Stewie 🐱 // be fly 🦋 like Quagmire 🦋 // roll like Joe 🎮
<https://twitter.com/wimdawgg2626/status/1602157104252477440>

These emojis are comparable to co-speech gestures, which occur during an utterance (Esipova 2019; Hunter 2019). Co-speech emojis and post-speech emojis may be difficult to distinguish from one another when the word/phrase modified by the emoji is utterance-final, such as the 🐱, 🦋, and 🎮 emojis in example (8). Gawne & McCulloch (2019) liken the semantics of post-speech emojis to that of co-speech gestures, so perhaps the two are not so different, semantically. Indeed, more semantic work will be needed to differentiate these cases. To me it seems that co-speech emojis iconically enrich syntactic constituents smaller than the level of an utterance, so the function of the emojis in (5) could be analyzed the same way as the activity emojis in (4).

I believe there is much interesting work to be done on co-speech emojis, but as this paper focuses on the morphology of the class of emojis described in the following section, I leave this for future work.

3.3. Pro-Speech Emojis

The third class of emojis, pro-speech emojis, are syntactically projected within a sentence. These are so named for pro-speech gestures, which are also iconic elements that are syntactically projected within a sentence (Schlenker 2018). They are incredibly frequent across CMC in a variety of languages², and are even commonly seen in some forms outside of electronic communication.

- (9) I ❤️ 🍺 [int: *love; beer*]³
<https://twitter.com/devil121710/status/1572384001121460224>

Several researchers have pointed out instances in which emojis appear as parts of speech within as sentence (Al-Rashdi 2015; Cohn et al 2018; Pierini 2021). Pierini (2021) calls these at-issue emojis, and compares them to existing analyses of pro-speech gestures. While some work suggests that these emojis merely replace words in a sentence, the facts I show in the remainder of this paper suggest otherwise. Take the following example from Pierini (2021).

² I do not have exact numbers on how many languages have speakers that use or accept pro-speech emojis, but I assume they are possible in any language whose speakers have access to emojis.

³ For each English example containing pro-speech emojis, I give an approximate translation in brackets of what the emoji means. Examples in sentences with more than one pro-speech emoji are separated by semicolons.

(10) She is the 🧨 [int: *bomb*]

(Pierini 2021)

In this example, if the emoji 🧨 is understood to be a replacement in the text for the character sequence <bomb> (which is itself a representation of the English word /bʌm/), then the emoji is not distinct from the actual English root /bʌm/, it is merely a unique way of representing the same lexical item that the text sequence <bomb> does.

The central argument of this paper, to be discussed and analyzed in detail in the remaining sections, is that emojis need not be associated with covert spoken words, which is something Schlenker (2017:38) claims of pro-speech gestures as well. I claim that pro-speech emojis in many cases represent distinct lexical items from the words to which they are thought to correspond, and are not words at all, but in fact free stems. I do this by demonstrating that, across a variety of languages, emojis can take affixes, which would not be expected if the emojis are merely replacing words that are already part of a sentence's numeration. Furthermore, these affixes often times do not correspond to the morphology of the word that is associated with an emoji. I then show that emojis can appear in cases where there is no clear equivalent in spoken language at all, and that emojis can under morphosyntactic processes such as lexicalization and grammaticalization.

The use of pro-speech emojis has several potential factors, including but not limited to abbreviation, iconic enrichment, or simply stylistic choice of the speaker. While language users' motivations in selecting a pro-speech emoji over a preexisting word are certainly an interesting area worthy of more exploration, this choice has no effect on the morphological generalizations I explore in the following sections⁴.

4. Morphology of Pro-Speech Emojis

In this section I introduce data showing emojis in places where lexical categories would appear in various languages, mostly English, Spanish, and German. All of the data, unless noted otherwise, come from Twitter. I collected these data using Twitter's search tool, which allows users to search for a specific token (in this case, an emoji) and filter by several different factors, the relevant ones here being language and date. Sometimes the search tool will be unable to distinguish between emojis, so often times searching for one emoji with a given affix or between a specific set of words will yield multiple different results. For example, a search "is ❤️ing lang:en" would return examples of pro-speech emojis with present-progressive verbal morphology following the word 'is' in English. A search that returns few results I take to be poorly attested and a search that returns no results I consider unattested and ungrammatical, and is thus preceded by an asterisk. Many of the English examples are further corroborated by the intuitions of native speakers who are proficient with emojis as a form of communication (myself included), and the examples from this paper in languages other than English are similarly confirmed by native speakers of those languages who are proficient with emojis.

⁴ When reviewing the data laid out in the following sections the reader may notice some key difference among some pro-speech emojis: some are more iconic than others while some seem to not be iconic at all, and some have a less obvious way to pronounce them than others do. These differences are not trivial and will be useful in further categorizing pro-speech emojis into at least two distinct categories, both of which conform to the morphosyntactic generalizations laid out in the paper, and will be discussed in Section 6.

A drawback of Twitter’s search tool is that it does not return specific numbers of results, so if a search yields a volume of results that is too high to manually count, then the exact number of results remains unknown. All of my examples – unless otherwise noted – are not isolated occurrences; rather, there are several pages worth of results. Anything with over 100 results I consider well-attested, and the examples I show are – unless otherwise noted – well-attested. Twitter does have an API tool for developers and researchers that can return more specific information such as the exact number of results and certain demographic information about users, but it this tool is difficult to access and even more difficult to use, so I leave it for future work.

In addition to showing emojis as syntactic constituents, I also show them combining with different kind of affixes. A discussion of what this means for the morphological categorization of these emojis will follow.

4.1. Inflectional Affixation

Pro-speech emojis appear where one would normally expect a word. Examples of these are numerous and easy to find in a multitude of languages and in several syntactic categories. Verbs in (11), nouns in (12), and adjectives in (13).

- (11) a. I ❤️ TEXAS! [int: *love*]
<https://twitter.com/JackManifoldTV/status/1514702804883021834>
 b. We 🤔 you Prime Minister... [int: *see*]
<https://twitter.com/Suebeloo1971/status/1516744745577553924>
 c. I’m gonna 💋 u on ur forehead u look 2 cute [int: *kiss*]
https://twitter.com/AnElfGirl_/status/1331319515263676420
 d. I need to 🛌 before I see the end of this game or I’ll be 😞 I missed it [int: *sleep, sad*]
<https://twitter.com/pjmdere/status/1259403629213224960>
- (12) a. My hair has gotten so long I feel like a 🧜‍♀️ [int: *mermaid*]
<https://twitter.com/alexisgarcia909/status/1514281876009091072>
 b. 🏠 is where the ❤️ is! [int: *home; heart*]
<https://twitter.com/GerryGanew/status/1516147151398576129>
 c. i play the 🎻 & the 🎸 [int: *violin; guitar*]
https://twitter.com/N0Y4_/status/1252837988779741186
 d. joon’s eating 🧊 cubes from his iced ☕... [int: *ice; coffee*]
https://twitter.com/RMPICS_twt/status/1334600431083909129
- (13) a. Pratik is always good at heart, he is a 😊 person [int: *happy, good*]
<https://twitter.com/itshpvarma/status/1488269483877683200>
 b. This has made me chuckle I needed this after a 💩 day! [int: *shit(ly)*]
<https://twitter.com/adampeter2011/status/1509556269048381441>
 c. Just glad my 🇺🇸 passport allows entry to New Zealand.... [int: *Australian*]
https://twitter.com/xray_nick/status/1469382259090284557
 d. Some 🌈 people were discriminated against at protest grounds [int: *gay, LGBT*]
<https://twitter.com/waaazap/status/1317916141763940357>

Examples such as these are, of course, not limited to English. See verbs (14) and nouns (15) in several languages below.

- (14) a. Я ❤️ минское море и мою подружку Леру (Russian)
 I love minsk-ADJ sea and my-ACC girlfriend-ACC Lera-ACC
 ‘I love the Minsk sea and my girlfriend Lera’
https://twitter.com/allo_mef/status/1515060138972962821
- b. Al rato vamos a 👁️ a militares dando clases en kinder (Spanish)
 to-the while go-1.pl to see to soldiers giving classes in kindergarten
 ‘Pretty soon we’ll see soldiers teaching Kindergarten classes’
<https://twitter.com/amadeo1980/status/1503486064266203143>
- c. Agora eu ✈️ pelo mundo (Brazilian Portuguese)
 now I fly for-the world
 ‘Now I’ll fly around the world’
<https://twitter.com/thammyrocha17/status/1124061997384052743>
- (15) a. Les 🐘 de la colère ! Laissez nous en paix vous humains ! (French)
 The elephant of the anger! leave-imp us in peace you humans !
 ‘Angry elephants ! Leave us in peace humans !’
<https://twitter.com/ElodieProutiere/status/1281004619112632325>
- b. Der ⚽ rollt! (German)
 the football rolls
 ‘The football rolls!’
<https://twitter.com/FCBayern/status/1274333767591776257>
- c. el pinchazo en el ❤️ cuando te dicen algo que temías (Spanish)
 the thorn-AUG in the heart when you say-3.pl something that fear-2.sg.ipfv
 ‘The thorn in the heart when someone tells you something were afraid to hear’
<https://twitter.com/saarabarroso10/status/1508434782719459331>

It is abundantly clear from Twitter data and from the judgments of language users who are proficient with emojis that emojis can represent elements that are projected in the syntax of a sentence, and this is evident in many languages in the world. But is it appropriate to simply say that emojis represent words? The following data show that emojis can take affixes, which suggests that emojis represent something smaller than a word.

4.1.1 English Inflectional Affixes

See emojis in English taking third person singular present agreement (16), past tense inflection (17), present progressive inflection (18), and participial inflection (19).

- (16) a. He ❤️s to 📖 [int: loves; read]
<https://twitter.com/cgoranson21/status/1514445141255372806>
- b. If in bed and she 👁️s his face while we do the nasty....I'm cool wit dat! [int: looks at]⁵
<https://twitter.com/bboybronx170/status/1498372492787961866>

⁵ Many speakers will add an apostrophe between an emoji and an affix. This seems to be nothing more than an arbitrary stylistic choice.

- c. But you must walk him often and he 🍌's his diaper a lot. [int: *poops*]
<https://twitter.com/BOBjustBOB2020/status/1500271892267110401>
- (17) a. My therapist 👻ed me so I took selfies in the parking lot [int: *ghosted*]
https://twitter.com/esq_ueer/status/1438573309474529282
 b. IMO we 👁ed better without him. [int: *looked*]
<https://twitter.com/canawupass/status/1439713360262615045>
 c. ...❤️d each other like brothers... [int: *loved*]
<https://twitter.com/Rossboylan1/status/1500537278112100353>
- (18) a. fuck disney plus im 🏴‍☠️ing [int: *pirating*]
<https://twitter.com/VAMPIREMANSION/status/1551747344558968833>
 b. likeeee the secondhand embarrassment is 💀ing meee [int: *killing*]
<https://twitter.com/innumando/status/1592854980176007175>
 c. ...you may Think they 🏃ing marathon to win gold... [int: *running*]
<https://twitter.com/ZigwanaSimon/status/1514220154179276800>
- (19) a. This was someone else's twt, but I would have ❤️d it had it been yours [int: *loved*]
<https://twitter.com/JPinNV/status/1479846200354299906>
 b. This world has 💋ed my soul with its pain, asking for its return in code [int: *kissed*]
<https://twitter.com/catsbi/status/1233358660908113920>
 c. We look forward to seeing what gets you 🔥ed up next. [int: *fired up*]
<https://twitter.com/MaverikLacrosse/status/1516466349266157571>

English also allows the pluralization of nominal emojis.

- (20) a. The way 🤍's act like it was so long ago when its just in their backyard [int: *whites*]
<https://twitter.com/DanielleMN30/status/1505580732323872770>
 b. Good luck to all our 🐬s swimming today [int: *dolphins*]
<https://twitter.com/DulwichDolphins/status/1515227917076340738>
 c. a couple of 🚬s smoking 🚬s [int: *fags; fags*]
<https://twitter.com/hailpaimin/status/1209546406379147269>

Comparative and superlative affixes also appear on English adjectival emojis.

- (21) a. Bitch say it to my face I'll drop u 🍌er than 4 o'clock [int: *deader*]
<https://twitter.com/Dxffyfy/status/1516834957422960644>
 b. This is the 🤍est scene ever created! I'm dying!!!! [int: *whitest*]
<https://twitter.com/DrVirgo1981/status/1479927717491621889>
 c. 📈er than the price of gas [int: *higher*]
<https://twitter.com/trujillojordan/status/924408241337044993>

The main generalization from these English data are clear: pro-speech emojis can take affixes. Before making more meaningful claims about this, though, it is necessary to look at data from other languages with different morphology. I use German and Spanish here.

4.1.2. German Inflectional Affixes

These processes are not only possible in English. Take the following German examples showing verbal affixation (22), pluralization of nouns (23), diminutive affixes on nouns (24), and comparative/superlative morphology on adjectives (25).

- (22) a. Und dann kommen erst mal 3 Tweets, die Du ge❤️t hast.
and then come.3pl {for now} 3 tweets that you PTCP-love-PST have.2sg.pst
'And then there's the three tweets that you've liked'
<https://twitter.com/DrStevieWer/status/923148032526442496>
- b. Mehr so vornerum ge👁️
more so in.front.of.around PTCP-eye.v
'More like eyeing around in front'
<https://twitter.com/SylviaKoopman10/status/1497561137369464837>
- c. Wir ■en alle!
we block-1.pl.pres all!
'We block everyone!'
<https://twitter.com/Usleon/status/1004673025973936128>
- (23) a. Vielen lieben Dank für die ☆e, liebe
many fond thank for the star-PL love
'Many thanks for the stars, love'
<https://twitter.com/peperonisahne/status/565079472173760512>
- b. Ich geb zu, ich habe gehamstert 🐱 alles für die 🐱n
I {admit} I have hoarded 🐱 everything for the cat-PL
'I admit, I've been hoarding – (I'd do) everything for the cats.'
<https://twitter.com/Menschin2015/status/1250448985862606852>
- c. Gewitter 🐐n, 🐱n, 🐷e, blaue Kartoffeln un Scheekönige. Biz jetzt super!
thunderstorm goat-pl, cat-pl, pig-pl, blue potatoes and snow.kings. Until now great
'A thunderstorm, goats, cats, pigs, blue potatoes and snow kings. So far all is great!'
<https://twitter.com/SoPaedGedoens/status/1000807383478784009>
- (24) a. Guten Morgen meine ❤️chen
good morning my dear-DIM
'Good morning my dear ones'
<https://twitter.com/Nina10109478/status/1078166969914937344>
- b. Wünsche dir heute 🎵🎶❤️ und immer ein lecker ☕chen
wish.1sg.prs you.dat today luck-music-love and always a tasty coffee-DIM
'Today, I wish you luck, music, love, and that you may always have nice coffee'
<https://twitter.com/kruemel6767/status/1156070352436568065>

c. Wer bis hierher gelesen hat: ich mag dich. Kriegst ein ☆chen
 who until here read.ptcp has: I like you.acc. get.2sg.prs a little-star-DIM
 ‘(To) those who have read all through here: I like you, you get a star’
https://twitter.com/Miss_Teached/status/1224032209717415937

(25) a. Ihr seid unsere ♥sten Menschen.
 y’all are our beloved-SPRL humans
 ‘You are our most beloved people’

<https://twitter.com/mariajlorenz/status/1201777426201751553>

b. ♥sten Dank liebe Alexandra, Selbiges wünsche ich dir auch
 heartfelt-SPRL thank dear Alexandra, the.same wish.1sg.prs I you.dat also
 ‘My most heartfelt thanks, dear Alexandra, I wish the same to you’

<https://twitter.com/hermann1959/status/1409221525727846406>

It seems, then, that German pro-speech emojis can take the same kinds of inflectional affixes as English emojis, including the diminutive suffix (which does not really exist in English). In the German examples we can see cases of more robust agreement systems than English involving verbal agreement markers and plural markers on nouns that change based on Case. This is not surprising since German has a much larger inventory of inflectional affixes than English does.

Despite differences in the number of available inflectional affixes, English and German share many similarities in their inflectional morphological systems. English inflectional morphology is always done to an uninflected base form that itself can stand alone as a word (or a ‘free form’ as described in Bloomfield (1933)). The same is true of German nouns and adjectives, but not necessarily of verbs (Kastovsky 1994). This means that for English inflectional stems and for German nominal and adjectival inflectional stems, there is an unmarked base form that also corresponds to a possible freestanding word. For German verbs, however, there is no unmarked base form. This means that German verbal stems need not correspond to a possible freestanding word.

Kastovsky (1994) describes the inflectional system of German nouns and adjectives and of English as a whole as word-based, since the inflectional stems can function as standalone words. He contrasts this with a stem-based system, such as the one for German verbs, in which the base form cannot function as a freestanding word. Before continuing with this discussion of morphological levels, it is necessary to look at data from another, non-Germanic language.

4.1.3. Spanish Inflectional Affixes

Pro-speech emojis in Spanish can be seen taking a variety of inflectional affixes as well. Plural suffixes are well-attested for nouns (26) and adjectives (27).

(26) a. Que los 🐶s sigan ladrando
 that the-masc.pl dog-pl continue-3.pl.sbjv stealing
 ‘May the dogs keep on stealing’

<https://twitter.com/AleBustamantee/status/1314014312864739330>

b. Si captas o saco las 🍏s y las 🍏s.
 if capture-2.sg or take-1.sg the-fem.pl pear-pl and the-fem.pl apple-pl
 ‘If you capture or I take out the pears and the apples’

https://twitter.com/Mx_Belle/status/1490023994027876359

c. Unos ☕s para el frío y sueño
 some coffee-pl for the cold and dream
 ‘Some coffees for the cold and tiredness’

<https://twitter.com/giulifoschi/status/729146304090247168>

(27) Podría comer chilaquiles todos los 🐶s días.
 can-1.sg.cond eat chilaquiles all-masc.pl the-masc.pl dog-pl days.
 ‘I could eat chilaquiles every fucking day’

<https://twitter.com/unforaneo/status/1378012687947325440>

The Spanish diminutive suffix (28) and augmentative suffix (29) can also be seen with pro-speech emoji nouns.

(28) a. Solito como un 🐶ito me dejarom
 alone-DIM like a dog-DIM me left-3.pl
 ‘They left me alone like a puppy dog’

<https://twitter.com/gregori52170557/status/1337561739135770625>

b. Alto finde, el ❤cito lleno de amorrr
 high weekend the-masc.sg heart-DIM full of love
 ‘Great weekend, my heart is full of love’

<https://twitter.com/Ayecaffaratti/status/1515842594575732736>

c. Fan del ☀cito de otoño-invierno
 fan of-the sun-DIM of autumn-winter
 ‘Fan of the little sun of autumn-winter’

https://twitter.com/bucco_marti/status/1399092698578100225

(29) a. Así me pone el 🐶☀azo
 so me put-3.sg the dog sun-AUG
 ‘The fucking sun makes me like this’

<https://twitter.com/DonkhodeHelio/status/1060604490376601600>

b. Otro ⚽AZO de Ángel Robles y ⚽ de Guillermo Martínez
 other goal-AUG of Angel Robles and goal of Guillermo Martínez
 ‘Another big goal for Angel Robles and a goal for Guillermo Martínez’

<https://twitter.com/ClubPueblaMX/status/1367901703362682880>

c. No pues gracias por el ✈azo monumental!
 no then thanks for the-masc.sg plane-AUG monumental
 ‘No thanks for the monumental disaster!’

<https://twitter.com/RockdeGallos/status/1426377757831860225>

According to Twitter and the intuitions of many native Spanish speakers, these are the only inflectional affixes with pro-speech emojis that are possible in Spanish. This notably leaves out

gender suffixes on nominals and verbal agreement/tense suffixes, which is important for this analysis.

Observe the following examples, which are confirmed ungrammatical by native Spanish speakers who are proficient emoji users. Additionally, Twitter searches for emojis followed by verbal phi-feature and tense agreement markers (and infinitival suffixes) and those followed by nominal/adjectival gender suffixes yield zero results.

- (30) a. *Yo ❤️ o a mi novio [int: *yo amo a mi novio*]
I love-1.sg to my boyfriend
'I love my boyfriend'
- b. *Quiero 👁️👁️ ar la tele [int: *quiero mirar la tele*]
want-1.sg watch-INF the TV
'I want to watch TV'
- c. *Nos 💋 ábamos [int: *nos besábamos*]
refl kiss-1.sg.ipfv
'We kissed each other'
- (31) a. *Todos los 🐶 os días [int: *todos los perros días*]
all the dog-m.pl days
'Every fucking day'
- b. *Toda la 🐶 a gente [int: *toda la perra gente*]
all the dog-f.sg people
'All of the fucking people'

With this additional data showing what is and is not possible with pro-speech emojis in Spanish, the task of coming up with a morphological theory of these emojis becomes much clearer. I make sense of these examples in the following subsection.

4.1.4. Emoji Stems

At this point it is clear that pro-speech emojis do not merely replace words in a sentence after the derivation has been completed: they represent something smaller than a word that actually is part of the morphosyntactic derivation of the sentence. Since emojis take affixes, they must represent something below the level of a word in a hierarchy of morphological levels that is something akin to the following (Lieber 1981; Selkirk 1982; Olsen 1986):

- (32) Words > stems > roots

I take 'word' here to mean any freestanding form of a given syntactic category with any acceptable number of inflectional and derivational affixes. Words are typically independently meaningful and often the target of syntactic operations such as movement. Stems, the level below words, are a lexically-typed base to which inflectional affixation is done. Stems may themselves contain affixes and their forms may or may not correspond with an acceptable freestanding words. Roots are bare elements containing zero affixes of any kind and, while possibly appearing only as a specific part of speech, they themselves are not specified for any syntactic category. There is much variation of these definitions in the literature (Kiparsky 2020;

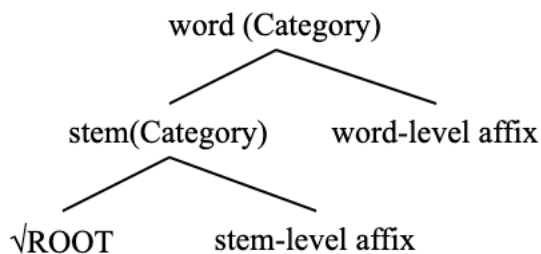
Londhal 2020). These notions – especially that of a stem – are very important to this analysis and will be explored in detail in this section and following sections.

Since roots contain no syntactic category, I believe that the predictions based in saying that pro-speech emojis are morphological roots would vastly overgenerate the amount of pro-speech emojis that we see and where we see them. We also would expect to see pro-speech emojis taking far more affixes than they actually do.

This leaves us with the stem. Stems are specified for a given syntactic category and are the target of affixation (Aronoff 1994). Notions of different types of morphological stems will be useful in determining what exactly pro-speech emojis are, morphosyntactically speaking. The idea of the level-ordering hypothesis in morphology (Bermudez-Otero 2018; Giegerich 1999, 2005; Kiparsky 2015; Siegel 1974) echoes what came up in Section 4.1.2. from Kastovsky (1994) about word-based and stem-based morphology by distinguishing between word-level and stem-level affixes.

To reiterate, all English inflectional affixes and all German nominal/adjectival inflectional affixes are word-based, while all German verbal affixes are stem-based. The terminology ‘word’ and ‘stem’ starts to get confusing as there are several slightly-competing definitions of these concepts floating around in the literature. Here I clarify that stem-level affixes promote categoryless roots to lexically-typed stems, and word-level affixes promote lexically-typed stems (which may or many not already be able to function as a freestanding word⁶) to inflected freestanding words. Therefore, the names ‘word-level’ and ‘stem-level’ refer to the *output* of the affixation, not the *input*. I now give an example of the representation of morphological structure assumed here on out. Roots are the lowest level that combine with an affix, that affix being a word-level affix. Intermediate stem levels combine with word-level affixes from which they are either promoted to words or become a stem for more word-level affixes. The highest node represents a freestanding word.

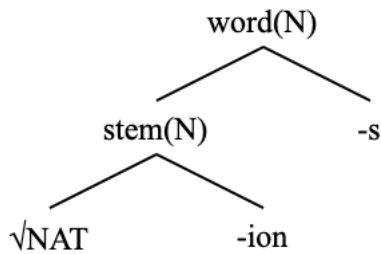
(33) Morphological structure of words



Before getting into pro-speech emoji examples, take the English word *nations* as an example of these levels. Here we have the root $\sqrt{\text{NAT}}$, the derivational stem-level suffix *-ion*, and the inflectional word-level suffix *-s*.

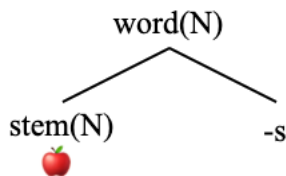
⁶ Here it is worth mentioning if zero morphology is necessary to promote a bare lexically-typed stem to a freestanding word, for example English singular nouns or first person present indicative verbs (Dahl & Fábregas 2018; Nida 1948). I believe that this is the case, though I do not think a full discussion of the theoretical implications of this is entirely relevant for this specific discussion.

(34) Structure of English *nations*



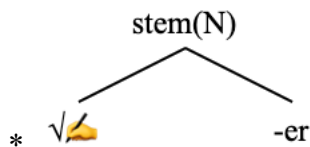
The root $\sqrt{\text{NAT}}$ must take some form of derivational morphology (here, the nominalizing suffix *-ion*) to be promoted to a lexically typed stem, otherwise it is not a candidate for inflectional affixation, hence the ungrammaticality of a form like **nats* with this root. Some English roots, such as $\sqrt{\text{OPT}}$ may be promoted to a lexically typed stem via a null derivational morpheme. With these generalizations in mind, let us take a grammatical example of a pro-speech emoji such as 🍏s, meaning *apples* [N, pl].

(35) Structure of English 🍏s



We can contrast this with an ungrammatical, unattested form such as *👨er, meaning *writer* [N, singular], the structure of which is given below⁷.

(36) Structure of English *👨er



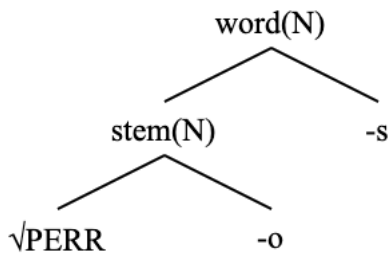
This contrast suggests that emojis are unavailable as roots and must instead be lexically-typed stems. As further evidence of this, I introduce the differences between two types of Spanish nominal suffixes: gender and number suffixes.

Gender suffixes are always closest to the root in Spanish, and number always follows gender. Take the Spanish example *perros* ‘dogs’, with the base form $\sqrt{\text{PERR}}$ followed by the masculine suffix *-o* and the plural suffix *-s*. The order ROOT-GEN-NUM is the only possible configuration

⁷ Note that the ungrammaticality of this example is not due to the choice of emoji. Indeed, Twitter data and native speaker judgments confirm that no emoji is possible in this position. Compare with other unattested forms *❤er (*lover*), *🚬er (*smoker*), *🍷er (*drinker*), etc. Note that the form *😘er (*kisser*) is unattested, except when in the compound 🍑😘er (*ass-kisser*), for which there are a few attested examples on Twitter. Several speakers find this compound notably more acceptable than the suffix with one emoji. I explain this in the following section.

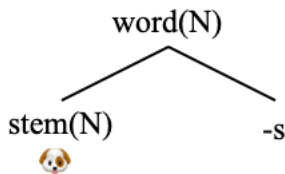
of these morphemes, and no other combination is possible. This means that gender is undeniably closer to the root than number, but how close exactly is a matter of some contention. The first possibility is that gender is an inflectional suffix that attaches to nominal stems (Picallo 1991; Alexiadou 2004). See *perros* with this internal structure, in which the root $\sqrt{\text{PERR}}$ must be an available candidate for affixation. The second option, which is that gender features are actually located on the nominalizing head, means that the gender suffix itself is what promotes the root to a nominal stem. This analysis is explored in detail in Kramer (2016). For our purposes, this means that the gender suffix is directly adjoined to the root, creating a nominal stem that can be inflected.

(37) Structure of Spanish *perros* with gender nominalizer



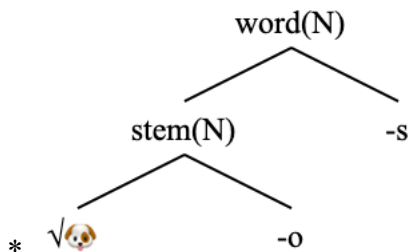
There is sufficient evidence to suggest emojis cannot be roots. With (37), the gender suffix is what turns the root into a nominal, as the gender feature is located on the nominalizing head in the analysis of Kramer (2016). As the stem, not the root, is the target of inflectional affixation, this means that the smallest possible stem in the word *perros* does not correspond with the root $\sqrt{\text{PERR}}$, but is the root plus the nominalizing head *-o*: *perro*. In other words, the gender suffix is a stem-level suffix, and the number suffix is a word-level suffix. So, the internal structure of 🐶s would be as follows.

(38) Structure of Spanish 🐶s



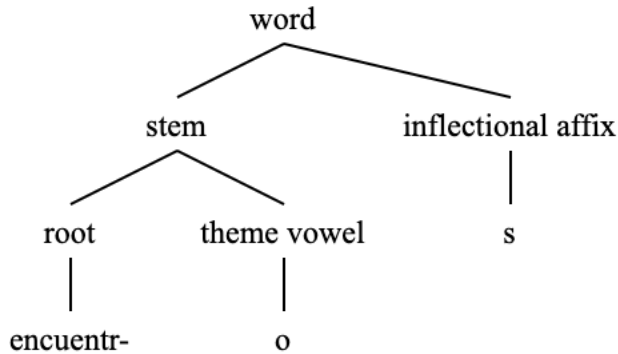
Since emojis are not available as roots, something like *🐶os is ruled out as an acceptable form.

(39) Structure of Spanish *🐶os



Assuming stems are what are stored in the language user’s lexicon (Zwicky 1992, Anderson 1992, Aronoff 1994), these bare root forms of Spanish nominals are not actually available as morphosyntactic primitives as they have no syntactic category. Stems, however, do contain this necessary morphosyntactic information. The gender suffixes of Spanish nouns and adjectives, also known as theme vowels, as they create a stem from the root, are the smallest unit in the lexicon of elements of these syntactic categories. Bermúdez-Otero (2013:5) gives the following tree, slightly modified for the purpose of this paper, for the noun *encuentros* [masc, pl].

(40) Structure of Spanish *encuentros*



In this structure, the stem is entered in lexicon, not anything lower. As this relates to pro-speech emojis, this means that they cannot appear as nouns or adjectives in Spanish as anything smaller than what is entered in the lexicon – anything smaller than a stem.

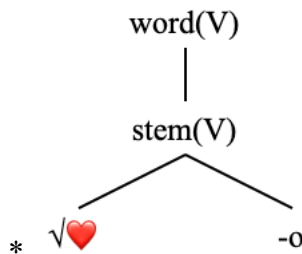
This paradigm also explains the inability of Spanish verbal emojis to take inflectional affixes, as these inflectional affixes also combine directly with the root. The differences between English and Spanish verbal emojis and this proposed explanation here are in line with preexisting observations about English and Spanish verbal morphology. We would not expect some primitive Spanish verbal root without affixes to exist in the lexicon. According to Lasnik (1995), French verbs are fully inflected in the lexicon of French speakers, so every possible inflected form of a verb is stored in the speaker’s lexicon, which is related to the fact that there are no bare forms of verbs, with even the infinitive forms having a suffix (the same is true of Spanish). Lasnik also argues that all English forms are bare in the lexicon, with verbal affixes being inserted in the syntax. Indeed, the same contrast between affixation with English and Spanish verbal emojis also exists with English and French, with an example such as (41) being incredibly well-attested online and confirmed acceptable by native speakers, but a form such as (42) being unattested and confirmed ungrammatical.

(41) Nous ❤️ Paris
 we love Paris
 ‘We love Paris’

(42) *Nous ❤️ons Paris
 we love-1.pl Paris
 ‘We love Paris’

This analysis of French verbal morphology can indeed be extended to Spanish, as Spanish verbs are also stored in the lexicon this way as inflected stems. For the Spanish verb *amar* (love), there is no root form in the lexicon *am-* without the infinitival suffix. Rather, there are lexical entries for every possible verb form including the infinitive and all inflectional combinations of person, number, and tense (*am-o*, *am-as*, *am-a*, *am-amos*, etc.). This unavailability of the bare form as a lexical stem means that the inflectional affixes are inseparable from the root on a lexical level. Since there is no bare verbal form available in Spanish, an emoji in the following position is not possible.

(43) Structure of Spanish *❤️o (*love-1.sg*)



Spanish verbal emojis cannot appear with affixes because there exists no level in the lexicon at which the inflected Spanish verbal stems are separable from their roots. This is different from English, where verbs are stored in the lexicon as bare stems, hence the grammaticality of (16-19). Spanish gender suffixes for nouns and adjectives, as well as Spanish verbal inflectional affixes, form part of the basic elements of the lexicon; they are stem-level affixes. All of the examples thus far demonstrate that emojis appear in positions of lexically-typed stems.

Perhaps the intuition that pro-speech emojis are words is not so far-off, an intuition reflected by the Oxford Word of The Year in 2015, which was the emoji 🤪. The truth of this intuition is that pro-speech only combine with word-level affixes, not stem-level affixes.

4.2. Derivational Affixation

Thus far we have only seen inflectional affixes. In English, all inflectional affixes are word-level affixes, but this clearly is not the case in a language such as Spanish. So, what happens in English with stem-level affixes? For this we must turn to derivational affixation.

It was my initial intuition that derivational affixes are not acceptable with emojis in English (or in any language, for that matter); however, that is not the case. Many English speakers accept these examples, and there is no shortage of them online. Examples including but not limited to adjective forming *-ly* (44), adjective forming *-y* (45) nominalizing *-ness* (46), and adjective forming *-ish* (47) are abundant.

(44) Have a ❤️ly Day! [int: lovely]

<https://twitter.com/iRedskinsFan/status/1516431824666382344>

(45) a. so 🧊y boys [int: icy]

<https://twitter.com/denimloveconnct/status/1374168048412676097>

- b. now no one can say I am being dramatic when I call it a 🤔y day!!!! [int: *shitty*]
<https://twitter.com/TaraMolinaTV/status/1428489881232232451>
- c. Just 🍎y. [int: *shitty*]
<https://twitter.com/ChrissMetz/status/1466622702903607304>
- (46) a. Happy New Year Krystal all the love and 😊ness to you [int: *happiness*]
<https://twitter.com/69darkone/status/1477160178286006272>
- b. Looking at this picture, just brings me great 😞ness [int: *sadness*]
<https://twitter.com/26humanbeing/status/1404677758228348929>
- c. Some days I think I've reached my peak 🏳️'ness. [int: *gayness*]
https://twitter.com/PDA_PhD/status/1416451272857763843
- (47) a. I real feel 🙄ish [int: *awkwardish?*]
https://twitter.com/marie_see/status/1545109718275850241
- b. This 🍌 is just getting 🙄ish. [int: *childish*]
<https://twitter.com/JRogers304/status/430766693187473408>

These examples are not to suggest that all derivational affixes are compatible with pro-speech emojis. In fact, most are not (recall the ungrammaticality of *🍌er in (36)). Take another unattested example that native English speakers proficient with emojis judge ungrammatical.

- (48) *🤔ity killed the cat [int: *Curiosity killed the cat*]

Once again, the choice of emoji is not what makes this sentence ungrammatical. Language users surely could make endless pedantic arguments about what emoji better conveys *curious* better among the thousands of possible emojis, but it is the position of the emoji that makes it ungrammatical. No emoji is attested as being modified by the nominalizing suffix *-ity*, whereas there are abundant examples of emojis being modified by *-ness*. There even are examples of emojis with *-ness* that have a similar interpretation to *curiousness*, which are also judged to be better than (48) by emoji-proficient native speakers.

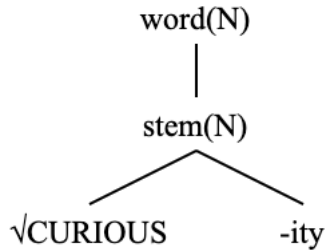
- (49) I feel like I remember him commenting on the 😏-ness of her outfit [int: *curiousness*]
<https://twitter.com/Cubsbane/status/1199749176667623431>

The difference between these cases is that *-ity* is a stem-level derivational affix, while *-ness* is a word-level derivational affix (Kiparsky 2020). This distinction between derivational affixes is also noted in Distributed Morphology (Embick & Marantz 2008; Marantz 2007). Here the difference is captured by the terms *inner derivation* and *outer derivation*. Inner derivation is the process of attaching derivational affixes to roots while outer derivation is the process of attaching derivational affixes to lexically-typed stems.

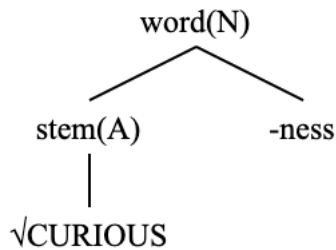
The difference between *curiosity* and *curiousness*, then, is that the *-ity* of *curiosity* attaches directly to the root $\sqrt{\text{CURIOS}}$, whereas *-ness* in *curiousness* attaches to the adjectival stem *curious*. This distinction is also visible in the phonological and orthographic effects that these processes have on their final products, as *-ity* changes the stress and spelling of the root, whereas

the root *curious* as it is pronounced in isolation is preserved in the word *curiousness*. Since pro-speech emojis cannot be roots, 🤪ness is acceptable as *curiousness*, but *🤪ity is not acceptable as *curiosity*. See the structures below.

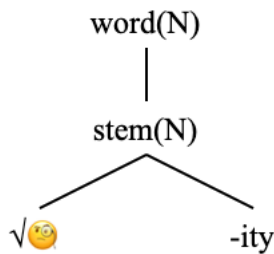
(50) Structure of English *curiosity*



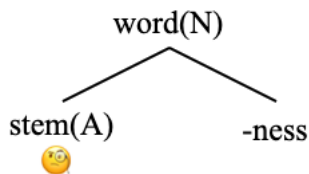
(51) Structure of English *curiousness*



(52) Structure of English *🤪ity



(53) Structure of English 🤪ness



It is the case that inner layer derivational affixes may not appear after a pro-speech emoji. This is completely expected under the assumption that emojis must be lexically-typed stems, not roots. This assumption is generalized as the following condition.

(54) *The Pro-Speech Emoji Condition*

Pro-speech emojis must enter a morphosyntactic derivation as a lexically-typed stem

I now turn to apparent exceptions in which emojis appear after inner-layer derivational affixes. Discussion of these data will reinforce the above condition

- (55) a. Le Président...is going to avoid the term 🇫🇮-ization from now [int: *Finlandization*]
<https://twitter.com/jwibo/status/1495744729299968004>
 b. same Montreal street before/after 🚶 i-ization [int: *pedestrianization*]
<https://twitter.com/spicygarage/status/1440086900191150081>
 c. TRUMP NK DE💣IZATION: FAILED [int: *denuclearization*]
<https://twitter.com/LaBelleEpoque/status/1089336625325400064>

Here, the inner layer derivational suffix *-ation* (which could be further decomposed into multiple derivational suffixes) appears after pro-speech emojis. This is not a problem for the theory above, though, and this is due to the fact that the emojis are immediately followed by the outer layer derivational affix *-ize*. This suffix can be observed without any other derivational affixes as well.

- (56) a. the whole trump administration literally...🔪ized it [int: *weaponized*]
<https://twitter.com/terrasapientiae/status/1420484100226494468>
 b. ...as symptoms of the end-stage disease of 🌐ized society [int: *atomized*]
<https://twitter.com/root2702/status/1152415656190652416>
 c. We're getting 🕌ized. [int: *Islamized*]
<https://twitter.com/75otingocni/status/1217571902211813376>

In the examples in (55), what happens is that nominal stem emojis are combined with the word-level affix *-ize*, to get a form such as 🇫🇮ize. This form then becomes its own root, which can be combined with other derivational affixes such *-ation*, so these examples are actually not an issue for the assertion that pro-speech emojis must be stems that combine with word-level affixes.

This reasoning also explains the potential acceptability of 🍑👄er (*ass-kisser*) when emojis cannot usually combine with the nominalizing suffix *-er*. First, the two stem emojis 🍑*ass*(N) and 👄*kiss*(V) form the verbal compound 🍑👄, which then gets promoted to a root that combines with the nominalizing suffix *-er*. As the emojis still enter the derivation as stems, they do not violate the generalization in (54). This is not the traditional way that a N+V-er compounds are understood to be formed in English, but there is no reason believe that pro-speech emojis must function the exact same way as their closest lexical counterparts. More examples of this are introduced in the following section.

4.3. Morphological Irregularities

Circling back to inflectional morphology, there are inflectional forms in English that are irregular which can be regularized in the presence of a pro-speech emoji. See the following examples.

- (57) a. counting 🐑s to fall asleep [int: *sheep*]
<https://twitter.com/togemuq/status/1292173903419174912>
 b. Lemme tell you bout these 3 blind 🐭s [int: *mice*]
https://twitter.com/ClinClin_Da_Kid/status/200805714720063488

- c. Finna search for some bigger 💎s for my 🦷s [int: *diamond;*, *teeth*]
<https://twitter.com/MikeyVJonzer/status/1357730531186601986>

Similarly, while the following examples are attested, there are no attested instances of of 🌵i or 🦀i, which are degraded.

- (58) a. whos idea was it to put 🌵's at the headboard of a bed tf [int: *cactuses*]
<https://twitter.com/hheydee/status/1516277845781192708>
b. We need 🦀s & 🦑es in that case! [int: *crabs*, *octopuses*]
<https://twitter.com/PetitionerDMin/status/1459841903323615232>

Irregular forms that involve ablaut or plural suffixes other than *-s* are not available for pro-speech emojis as the emojis themselves are the stem. In other words, the stem has no vowels to be ablauted. For the irregular plural suffixes such as *cacti*, the derivation would involve stem-level affixation to the root as these forms come from Latin, which forms plurals via stem-level derivation, which we have already established that emojis cannot do.

4.4. Interim Summary

So far I have demonstrated, among the ways in which emojis are used in language, there is a distinct category of emojis functioning as something that appears at first to be words, called pro-speech emojis. Upon further examination of data from Twitter as well as the judgments of native speakers in English, German, and Spanish, it becomes clear that pro-speech emojis felicitously and regularly are modified by both inflectional and derivational affixes. Analysis of the types of affixation possible across the three aforementioned languages as well as of current morphological theory strongly suggests that pro-speech emojis are only possible as stems of a specified syntactic category that take word-level affixes, which in English includes inflectional affixes and outer layer derivational affixes. Evidence from the regular inflectional morphology of pro-speech emojis whose closest lexical equivalent inflects irregularly suggests that pro-speech emojis are totally distinct from the words to which they are thought to correspond, which will be explored further on.

5. Pro-Speech Emojis and Language Change

As morphosyntactic elements which are part of the derivation of words and sentences, it should stand to reason that pro-speech emojis are sensitive to processes of morphological change, which is exactly what we see. This means undergoing category changes, being borrowed between languages, novel forms being innovated, and the like. In this section I discuss pro-speech emojis undergoing morphological processes of language change, namely lexicalization (Section 5.1.) and grammaticalization (Section 5.2.).

5.1. Lexicalization

In this section I discuss cases of emojis undergoing lexicalization. I identify three different types of lexicalization of emojis: expressive emojis becoming pro-speech emojis, emojis undergoing a category change, and pro-speech emojis being borrowed into a new language. I also compare these to processes of lexicalizing abbreviations.

5.1.1. Expressive to Pro-Speech Emojis

It is quite common to see expressive emojis functioning as verbs.

- (59) a. Well you can see me 🙄ing but yeahhh i signed up for this. [int: *facepalming?*]
<https://twitter.com/lordbongjae/status/1273951705550295042>
b. my brother caught me 🙄ing [int: *just standing there*]
<https://twitter.com/D00MB0YSEASON/status/1336839638259011586>

To focus on a specific example, take 🙄 as an expressive emoji.

- (60) a. He flew 40 hours from the UK to take me on our first date 🙄
<https://twitter.com/missbigtipper/status/1538677795878240256>
b. friendly reminder that Groggu got Din a father's day gift 🙄
<https://twitter.com/spoiler4you/status/1538731192207679490>
c. Pisces moon got me wanting some love & affection 🙄
<https://twitter.com/LoveOfDivine/status/1538667548849213441>

Now, observe the same emoji with verbal morphology. Here I leave out a translation of the emoji in brackets in the examples as I discuss it in the text.

- (61) a. Don't 🙄 at me, you little bottom
<https://twitter.com/trixiemattel/status/1326930133643526144>
b. never in my life have i 🙄ed at a girl like this
<https://twitter.com/chucken99/status/1407373890041483268>
c. can we all just take a second to 🙄 at the show's twitter bio
<https://www.tumblr.com/blog/view/roguebebe/629565055095226368>
d. are u 🙄ing other bitches
<https://twitter.com/fuckingtyrell/status/1226958016953667590>
e. I'm a top and I 🙄 all the time stop this slander
<https://twitter.com/bunnysnack/status/1244099514387980293>

What exactly this emoji means is hard to formalize. Many compare it to the expression 'uwu', but that is even more meaningless to those who are outside of communities who say that. Undoubtedly, 🙄 is highly expressive as it clearly resembles a human facial expression. *Emojipedia* describes as a "face with furrowed eyebrows, a small frown, and large, 'puppy dog' eyes, as if begging or pleading. May also represent adoration or feeling touched by a loving gesture." The website also mentions that it can be "somewhat suggestive" and even "horny" (Burge 2020). As a verb, then, 🙄 means something along the lines of 'to make the facial expression 🙄', with all its possible interpretations.

What we are seeing here is an expressive emoji being used as a verb. This is, in some sense, a category change, but not really. It is not appropriate to think of expressive emojis as having a syntactic category since they are post-speech elements (Grosz et al 2021), but anyhow this expressive element is being lexicalized as a verb.

Also interesting is that there is no clear spoken English equivalent of this verb. This is unlike the following cases, where the emojis could be compared to ‘smile(v)’ and ‘laugh(v)’, respectively.

- (62) a. This made me 😊 saw a 🐰 going nom nom while I was watching tv [int: *smile; rabbit*]
<https://twitter.com/Yolanda24464357/status/1542314126726336514>
 b. over here 😂ing at the Yankees [int: *laughing*]
<https://twitter.com/RealRidgeyO/status/1503205343282667522>

This is independent evidence that this emoji does not replace preexisting words, as there is no word that 😊 could replace. The data in (60-62), then, are to say that the expressive emoji has undergone lexicalization to a verb.

5.1.2. Category Changes

In Brazil in the spring of 2022 a video of a lobster went viral online. In this video, the lobster can be seen walking up to and jumping in a pot of boiling oil⁸. This video quickly became a meme that was shared in video and GIF form across social media. In the wake of this, the use of the emoji 🦞 shown below appeared.

- (63) a. vou me 🦞 hj 13h
 go-1.sg self kill today 13-hours
 ‘I’m going to kill myself today at 1:00pm’
https://twitter.com/ibizatrance_/status/1529854246690103296
 b. vc se 🦞
 you self kill
 ‘Kill yourself’
https://twitter.com/lucas_kapocz/status/1527292927847833600
 c. Começando o feriadão com cólica vou me 🦞
 beginning the holiday with colic go-1.sg self kill
 ‘Starting the holiday with colic I’m going to kill myself’
https://twitter.com/jp_90skid/status/1537195620037017606

Before this, the lobster emoji can be seen consistently as a pro-speech emoji of the category Noun in Brazilian Portuguese tweets.

- (64) a. Não vai mais comer 🦞 tão cedo
 No go-3.sg more eat lobster so soon
 ‘Don’t go eating eating lobster so soon’
<https://twitter.com/MascarenhasT/status/1508813227987292166>
 b. Vontade de comer 🦞
 wish of eat lobster
 ‘Feel like eating lobster’
<https://twitter.com/gfelipeceliipe/status/1267225810420801538>

⁸ The video itself can be viewed here: <https://www.youtube.com/watch?v=G2a9c240WVM>. A discussion of the virality of the meme can be viewed here: https://www.youtube.com/watch?v=Kn_MNGgxT6Q.

- c. Só vou parar quando acabar a 🦞 e 🍷
 only go-1.sg stop when finish the lobster and wine
 ‘I’m only going to stop when the lobster and wine runs out’
<https://twitter.com/RighettiSR/status/1273147151401521152>

In this case the pro-speech emoji 🦞 is undergoing a clear category change from Noun to Verb. Rather than creating a new lexical item from an expressive element, this is a reassignment of syntactic category to a preexisting emoji. One informant⁹ even states that they would pronounce the verbal emoji 🦞 as *lagostar*, which is the nominal form *lagosta* plus the verbal infinitival suffix *-ar*, which is what would be expected of a verb derived from a noun in Brazilian Portuguese. Interestingly, though, the meaning is completely unpredictable from the form, quite unlike a case such as *fish* (N) and *fish* (V) in English. Here the meaning is dependent on the shared understanding of the meme.

5.1.3. Borrowing of Pro-Speech Emojis

One of the main processes of lexicalization is borrowing from other languages (Hilpert 2019). This can happen with emojis as well. While it may seem like the same pro-speech emoji appearing in different languages is just an instance of two different lexical items in the same language (*el 🐏* in Spanish and ‘the 🐏’ in English, for example), this is not always the case. There are verifiable instances of pro-speech emojis being borrowed from one language to another.

Take the acronym G.O.A.T., which stands for ‘greatest of all time’ in English and is a very popular superlative on Twitter, especially used to describe athletes (Curtis 2017). Due to this acronym’s orthographic and phonological similarity to the noun ‘goat’, the emoji 🐏 is often used as a superlative as well. See the following example.

- (65) He is the 🐏 but I need to see the batfleck in his own movie
<https://twitter.com/Harlem917324771/status/1525716454523383808>

This superlative has been borrowed into Spanish as well, and is very commonly used by Spanish speakers on Twitter, also especially when it comes to athletes.

- (66) a. Garnacho sabe quien es el 🐏
 Garnacho knows who is the-masc.sg 🐏
 ‘Garnacho knows who the G.O.A.T. is’
<https://twitter.com/Unitedneta3/status/1524741444765896704>
- b. El 🐏 y Messi a su lado.
 the 🐏 and Messi to his side
 ‘The G.O.A.T. and Messi by his side’
https://twitter.com/IvanBlanco_26/status/1506694002170634243

⁹ This informant is also the person who explained to me the cultural background of the lobster suicide meme.

The masculine determiner in these examples is important to note because the word for goat in Spanish is feminine (*la cabra*), so it is not the case that this is an instance of a pro-speech emoji that corresponds to the Spanish word for goat being used in a novel way. Spanish pro-speech emojis typically contain grammatical gender information while not needing an overt gender suffix, which makes sense considering the nature of gender morphology in Spanish, and this gender typically corresponds to the gender of the emoji's closest synonym. While it may be possible that the emoji is being used in a way closer to a word such as *cabrón* 'bastard', which is masculine, the sentiment of the Spanish tweets containing 🐐 is almost entirely positive and is used in the same sports contexts and it is in English.

5.1.4. Abbreviations

The lexicalization processes seen above with emojis are not new in any way, which is perhaps why they are notable. They are immediately recognizable as similar to lexicalization of other elements, especially those specific to – or at least common in – CMC.

Abbreviations in CMC such as *lol* 'laugh out loud' or *lmao* 'laugh my ass off' represent VPs, but they are able to be lexicalized as verbs that can take objects (67), or even nouns (68). (67b) is especially interesting when considering that the possessive pronoun 'my' is clearly no longer part of the composition of the verb.

- (67) a. im on harry styles baldtok im lmaoing
<https://twitter.com/SYSCAbout/status/1546400157096628224>
 b. my gf just lmao'd at a non-funny tweet.
<https://twitter.com/Chrsboyy/status/1485113799883771905>
- (68) a. Twitter should ban Elon Musk just for the lols
<https://twitter.com/themaxburns/status/1545524331836542983>
 b. Thank you for a mighty LOL. That thread is gold.
<https://twitter.com/MaureenRogers49/status/1484933478802247683>

This is, of course, not specific to CMC, though. Similar processes lead to apparent redundancies such as 'ATM machine', 'PIN number', 'NYI institute', 'BIPoC people', and so forth.

Abbreviations also undergo borrowing between languages. Take the following examples of a French abbreviation in English, an English abbreviation in Spanish, and an English abbreviation in Russian.

- (69) So many ppl have already RSVP'd and I sent it 2 hours ago.
https://twitter.com/JOyfulneSS_/status/1545902209128448000
- (70) Pero es un LOL mayúsculo
 but is a LOL majuscule
 'But it's a capital LOL'
<https://twitter.com/Reisdeg/status/1536663646193041409>

- (71) имхо твиттер больше всего хейтят..
 imho twitter more everything-gen hate-prs.pl..
 ‘IMHO Twitter is the most hated’
<https://twitter.com/dysforiac/status/1546590702703976449>

The interpretation of these abbreviations does not depend on understanding what it stands for in its original language. For example, English speakers regularly say ‘RSVP’ without necessarily knowing what it stands for in French. This is clearly similar to the 🐶 example in English/Spanish.

5.2. Grammaticalization

If emojis can be lexicalized, surely they can undergo grammaticalization as well. This would entail an emoji going from a lexical element to a functional element (Campbell & Janda 2001; Roberts & Roussou 1999). Since emojis can indeed be lexical items, such a thing should be possible, and indeed it is attested. Take the following examples.

- (72) a. also we going ✈️ doctor tomorrow
<https://twitter.com/shawnobrjen/status/1074898134357237761>
 b. Going ✈️ Nevada to count these fucking ballots
<https://twitter.com/knotkate/status/1324545992897814535>
 c. going ✈️ supreme court ⚖️ y’all need anything?
<https://twitter.com/PledgeCarlee/status/1325145420155248646>
 d. IM GOING ✈️ MED SCHOOL
<https://twitter.com/Jwoo/status/1260479782816538624>
 e. going ✈️ doctors for this mf migraine
<https://twitter.com/fauxgemini/status/1392203437878628353>

Here the airplane emoji corresponds to the word ‘to’, indicating directionality and introducing the DPs that follow them. This emoji says nothing about the modality of travel, as in the emoji in (72a) does not contribute the meaning ‘we going to the doctor tomorrow *by plane*’. We also see the emoji in places where we would not expect the preposition to in spoken English.

- (73) a. going ✈️ vacation next week, but i’m ugly with nothing packed
<https://twitter.com/mashgravity/status/1134433810249109504>
 b. going ✈️ places
https://twitter.com/fanie_nani/status/1392144620478509058
 c. going ✈️ home for Christmas
<https://twitter.com/Zack3O12/status/1076028923857854464>
 d. we going ✈️ someplace to eat before the show starts
<https://twitter.com/guidemeback/status/1142544192239017984>

The spoken English preposition in (70a) would be ‘on’, and the words ‘places’, ‘home’, and ‘someplace’ in these examples would be introduced by a null preposition in spoken English and are ungrammatical with an overt preposition (Collins 2007; Schoenmakers & Storment 2021).

This construction is very common on the internet, even having its own Urban Dictionary entry. Grammaticalization creating prepositions is nothing new: ‘behind’, ‘betwixt’, ‘before’, ‘between’, ‘in front of’, just to name a few in English. If emojis are in the lexicon, there is no reason that they could not be grammaticalized as well. The connection between the iconicity of the airplane and a directional preposition is not hard to see. It is far more intuitive than using any other emoji, say 🛩️, as a directional preposition.

Another interesting example of a pro-speech emoji belonging to functional category is the emoji 👉 as a second person singular pronoun.

- (74) a. I’m looking at 👉
<https://twitter.com/nahhhman/status/1566623058336563200>
b. It’s simple... If 👉 engage with me, I’ll engage with 👉
<https://twitter.com/PrattyNFT/status/1583181509472620544>
c. Don’t worry about what others are doing. Worry about what 👉 are doing.
https://twitter.com/mythic_meta/status/1587425440959979525

This example’s iconicity is pretty straightforward, as the interpretation of the emoji comes from the fact that it is pointing at the addressee (the reader). This is in fact the same symbol as the sign for the second person singular pronoun in ASL. Once again we have an example of an expressive symbol becoming a pro-speech emoji, only now it is a pronoun.

Pro-speech emojis, then, can belong to functional categories as well as lexical ones. This is the only example of such a case that I am able to find, which should not be shocking given that at this point emojis are only eleven years old, and these language change processes take time. In the future we should expect to see more cases of grammaticalization of pro-speech emojis.

6. A Note on Classification

At this point it is crucial to ask not only what pro-speech emojis represent morphosyntactically, but what the nature of the difference is between pro-speech emojis and regular English *orthography* (or speech, for that matter). The way I see it, there are four options: (a) there is no perceived difference between a pro-speech emoji and other words; (b) the difference is one of modality, and the transition between pro-speech emojis and other words is parallel to bimodal code-switching in signed and spoken language multilinguals; (c) the difference is still one of modality, but the transition from written words to pro-speech emojis back to written words is parallel to that of using a pro-speech gesture in connected speech; and (d) the difference is purely orthographic, with pro-speech emojis representing some kind of enhanced orthography, comparable to the difference between Japanese hiragana and katakana.

It is safe to say that option (a) is not valid, since there is a body of work showing a difference in the way that emojis are processed from surrounding text (Cohn et al 2018; Barach et al 2021). The comparison to bimodal code-switching also feels lacking, as pro-speech emojis can do things that we do not see at the border of spoken/signed modalities, such as taking affixation of a different modality. There is also no evidence to suggest that pro-speech emojis are part of anything that could be considered a different language from the surrounding text, and it intuitively is inappropriate to assert that monolingual emoji users are actually bilingual.

It is for these reasons that (c) or (d) seems most likely. As we see in languages that regularly make use of distinct orthographic systems with frequent switches between them, emojis serve distinct morphosyntactic and discourse functions. In this case, pro-speech emojis would be an enhanced expressive orthography for an extended or enhanced lexicon, although there are some observable differences within the category of pro-speech emojis that must be identified and discussed. Based on the data available in this paper and from Twitter as a whole, two separate patterns emerge simultaneously. In one, pro-speech emojis function as a kind of code for preexisting words, in which their interpretation relies on an association between the form of the emoji and a distinct lexical item that is available to the speaker/reader (which is, of course, not to say that the emojis actually replace a similar word in any point of the derivation of a given sentence in which they appear). In the other, the pro-speech emoji may be interpreted fully – or mostly – iconically, without making reference to any other lexical item to be understood. I dub these *code emojis* and *primitive emojis*, respectively. In the remainder of this section I discuss details and examples of these two new categories, as well as some basic diagnostics for differentiating the two. Up until now this paper has not distinguished between these two categories as the morphosyntactic generalizations of pro-speech emojis hold up of both categories, but all of the aforementioned data can be sorted as either a code emoji or a lexical emoji. It is important to note that the morphosyntactic generalizations laid out in this paper still hold of both code emojis and primitive emojis.

6.1. Code Emojis

As stated above, the interpretation of code emojis relies on the language user's ability to make an association between the pro-speech emoji and a related lexical item, and retrieve the meaning that way. As such, code emojis do not iconically represent what they denote. See the following examples.

(75) 👻'ed him bc I'm tryna get in the Halloween spirit [int: *ghosted*]
https://twitter.com/Dede_1204/status/1576619360336027649

(76) anyone else think 🏴‍☠️ing movies is a good thing? down with capitalism... [int: *pirating*]
<https://twitter.com/bBlueberrybrow/status/1288579081299202050>

Here the emojis are not being interpreted iconically and are instead a reference to the idiomatic meanings of the English verbs 'ghost' and 'pirate'. A test to show that these emojis are not iconic in their interpretation is that they are not available to use in this way in languages that lack similar lexical items. For example, in Chinese, the equivalent to the verb 'ghost' in English, meaning to cut off communication (often of a romantic nature) with someone suddenly and without explanation, would be 鸽 'to pigeon'. See the following example from an informant.

(77) 他🐦了我
 he pigeon-perf me
 'He pigeoned (ghosted) me'

Note that swapping the pro-speech emojis in (72) and (73) would yield uninterpretable results in both English and Chinese, as their interpretations rely upon the availability of the interpretations of these lexical items in each language.

A similar point can be demonstrated language-internally, as well. Take example (20c) of this paper, repeated below.

(78) a couple of 🚬s smoking 🚬s [int: *fags*; *fags*]
<https://twitter.com/hailpaimin/status/1209546406379147269>

The original Tweet here contains a picture of two gay men in the bathtub together smoking cigarettes. The appropriate reading of the sentence, then, is ‘a couple of fags smoking fags’. Preliminary judgment tasks confirm that this particular sentence is very difficult for people to parse who are not aware that some gay men use the word ‘fag’ as an in-group self-identifier, or that the word *fag* in British English means *cigarette*. This means that the interpretation of the first 🚬 emoji relies on the user’s knowledge of first the British English word and second the repurposed use of this slur. This is especially crucial to the appropriate interpretation of this Tweet since the second instance of the 🚬 emoji is simply an iconic representation of a cigarette.

Another example from queer subculture is the emoji 🏳️, which is often used by queer people online to refer to the camp aesthetic. An example of this is shown below.

(79) I’m tired of pretending to not like riverdale, it’s 🏳️ [int: *camp*]
<https://twitter.com/OctAirie/status/1507865498138161155>

This is a clear instance of the 🏳️ emoji referring to ironic bad taste, as opposed to the many instances of people who are not part of subcultures that regularly discuss the camp aesthetic using the 🏳️ emoji to refer to actual tents and campsites.

As codes for words, we would expect these emojis to always have an easily available pronunciation, so if a person were asked to read a sentence containing a code emoji aloud, they would not struggle to find an equivalent spoken word for the emoji, which intuitively seems to be the case, though experimental results to demonstrate this will be valuable. We would also expect code emojis to be *less* likely to display the morphological regularization seen in (52-53) of this paper, as there could be phonology associated with code emojis available to the speaker that is not available with lexical emojis. This could also explain the alternation between Spanish diminutive suffixes *-ito* and *-cito* in (28), which is typically understood to be a morphophonological alternation.

More work is definitely needed with pro-speech code emojis, especially to establish more concretely the semantic connection between these emojis and the concepts they are related to, perhaps by some means of semantic motivation.

6.2. Primitive Emojis

Primitive emojis do not make reference to any preexisting English words. Their interpretation is iconic, even if their interpretation is that of a highly specific object/concept. Often times there is no salient phonetic form of a primitive emoji, which can lead to confusion when a person is asked to read aloud a sentence containing a primitive emoji. As stems without phonetic form that do not even make reference to any elements with phonology, these are the emojis with which we would predict to see the most morphological regularization, meaning we would not expect to see anything like irregular pluralization or the use of the suppletive determiner in English. Despite being potentially more difficult to pronounce than code emojis, it is generally not the case that primitive emojis are more difficult to understand. In fact, due to their iconicity, they are generally the ones that are more straightforward to interpret.

This is not to say, however, that every primitive emoji is unpronounceable. Take the following example, which is notably very easy to interpret.

(80) I ❤️ my 🐱s [int: *love*; *cats*]

Few – if any – people who have seen emojis before would struggle to read this sentence out loud, and it is sufficient to say that this is because the semantic denotations of the noun 🐱 and the verb ❤️ are almost identical to those of the noun *cat* and the verb *love*. Although, interestingly, the primitive emoji verb ❤️ has gained a new pronunciation, which *heart*. Preliminary results would suggest that a similar number of people would read (77) as ‘*I heart my cats*’ to those who would read it as ‘*I love my cats*’. This phenomenon of an emoji gaining a new pronunciation is also something that would only be expected with primitive emojis, not code emojis.

Cases in which there is an emoji with no obvious phonetic form and a clear iconic interpretation, such as 😏 discussed in Section 5.1., are unambiguously primitive emojis. It cannot be the case that this emoji’s interpretation relies on the knowledge of another lexical item because there is no other lexical item in English with a meaning that can be roughly represented by the following.

(81) $\exists e$ [making the 😏 facial expression(e) & Agent(e, x) & Goal (e,y)]

It is also clear how this interpretation when applied to the emoji 😏 is purely iconic. Even cases such as the verbal 🇧🇷 in Brazilian Portuguese have an iconic interpretation, even if that interpretation necessitates knowing about the viral video on which it is based.

It is important to return to the grammaticalized ✈️ emoji from Section 5.2. here. At this point, it may seem like the ✈️ emoji as a functional element fits into neither of the categories designated for pro-speech emojis. It does not work as a code emoji because there really is no clear phonological equivalent for the emoji, and it does not work as a primitive emoji under the stipulation that it is not iconic. Planes travel unidirectionally from one location (source) to another (goal). The orientation of the plane emoji is such that it is pointing to the right, which means in a standard English sentence (and indeed in all of the examples given) is pointing toward the goal. As such, its iconicity as a directional preposition that introduces a goal is not a tough sell.

Additionally, it could be the case that this emoji originally denoted specifically travel by plane, and overtime became a more general functional element, which would very much support a semantic bleaching account of grammaticalization (Sweetser 1988). This is also supported by more recently-attested data in which very similar elements do have a more obvious iconic component, indicated modality of travel or otherwise.

(82) a. Going 🛫 Gilded Age

<https://twitter.com/itgetsbedder/status/1501015675715964935>

b. going 🌿 river day

<https://twitter.com/MilkmanNick/status/989956661832957952>

It looks, then, as if pro-speech emojis that have undergone lexicalization or grammaticalization are more often primitive emojis, but this is not always the case, as the use of 🐎 as the abbreviation G.O.A.T. in English or Spanish is not iconic.

With this distinction, the use of primitive emojis is most accurately characterized by being parallel to using a pro-speech gesture in connected speech, while code emojis are similar to an enhanced expressive orthography for elements associated with lexical items, which perhaps opens up the possibility of analyzing the semantics of code emojis as being similar to co-speech gestures (Esipova 2019).

A major question raised by the dichotomy of code emojis and primitive emojis is whether or not primitive emojis actually exist as entries in a person's lexicon. Obviously, it is not ideal to posit that every primitive emoji noun or verb is a distinct entry in a person's lexicon, just as it is not ideal to posit that any possible pro-speech gesture in spoken language is represented in someone's lexicon. It may be advantageous, however, to state that some primitive emojis with conventionalized uses (i.e. 😬 as a verb, ✈️ as a preposition) do exist as entries in the lexicons of those people who regularly use them, especially for those cases like ✈️ where the iconicity has become less salient in some way. Just as words, these emojis in and of themselves are not designated as code or primitive pro-speech emojis, or even pro-speech emojis at all for that matter. It depends on the contexts in which they appear. As such, the ✈️ as an activity emoji, or even as a primitive emoji noun would not trigger a proposed lexical entry for the ✈️ preposition.

7. Conclusion

In this paper I demonstrate that there is a distinct class of emojis – pro-speech emojis – function as independent morphological stems that take word-level affixes, which, in English, include all inflectional affixes and outer-layer derivational affixes. Data supporting these facts are *incredibly* numerous, well-attested, and confirmed by native speaker grammaticality judgments. The existence of some of these pro-speech emojis can also be analyzed as a result of lexicalization and grammaticalization. Pro-speech emojis are attested as belonging to the categories Noun, Adjective, Verb, Preposition, and Pronoun, although in this framework theoretically any syntactic category could be assigned to a pro-speech emoji. Within pro-speech emojis, two potential classes have been identified: code emojis, whose interpretations rely on the meaning of

another lexical item, and primitive emojis, whose meaning is iconic. There are several diagnostics to differentiate these two classes, including iconicity and possible phonetic form. This work builds on ideas introduced in the super-semantics program, namely that symbolic elements can be analyzed in a formal linguistic system (Greenberg 2021; Schlenker 2019). My work here demonstrates that symbolic elements actually can be part of the linguistic system themselves, where they conform to regular morphosyntactic principles. This makes a much-needed connection with work on the semantics of emojis (Grosz, Kaiser & Pierini 2021; Grosz et al 2022; Maier 2020; Pasternak & Tieu 2020; Pierini 2021) to morphology, and by extension syntax (Collins & Kayne 2021). This opens up the possibility of looking more closely at the morphology and syntax of other pro-speech elements such as gestures or even music (Migotti & Guerrini 2022).

Moving forward with this analysis, it will be crucial to continue to gather more data on this topic, ideally from languages not mentioned in this paper and with sociolinguistic variables such as age and location. A large-project involving statistical analysis of data using Twitter's API tool or something similar will be massively helpful.

8. References

Al Rashdi, Fathiya. (2018). Functions of emojis in WhatsApp interaction among Omanis. *Discourse, Context & Media*. 26.

Alexiadou, A. (2004). Inflection class, gender and DP-internal structure. In G. Müller, L. Gunkel, & G. 520 Zifonun, eds. *Explorations in Nominal Inflection*. Berlin: Mouton. 21-50.

Alshenqeeti, H. (2016). Are Emojis Creating a New or Old Visual Language for New Generations? A Socio-semiotic Study. *Advances in Language and Literary Studies*, 7, 56-69.

Anderson, S. R. (1992). *A-morphous morphology*. Cambridge, UK: Cambridge University Press

Arafah, Burhanuddin & Muhammad, Hasyim. (2020). Linguistic functions of emoji in social media communication. *Opcion*. 35. 20.

Aronoff, Mark. (1994). *Morphology by Itself: Stems and Inflectional Classes*. MIT Press

Bai, Q., Dan, Q., Mu, Z., & Yang, M. (2019). A Systematic Review of Emoji: Current Research and Future Perspectives. *Frontiers in Psychology*, 10.

Barach, Eliza & Feldman, Laurie & Sheridan, Heather. (2021). Are emojis processed like words?: Eye movements reveal the time course of semantic processing for emoji-fied text. *Psychonomic Bulletin & Review*. 28.

Bermúdez-Otero, R. (2013). The Spanish lexicon stores stems with theme vowels, not roots with inflectional features. *Probus*, 25(3–103).

- Bermúdez-Otero, R. (2018). Stratal phonology. In S. J. Hannahs & A. Bosch (Eds.), *The Routledge handbook of phonological theory*. 100-134. London, UK. Routledge.
- Bloomfield, Leonard. (1933). *Language*. New York: Holt.
- Burge, Jeremy. (2020). A New King: Pleading Face. *Emojipedia*. Blog
- Campbell, Lyle & Janda, Richard. (2000). Introduction: Conceptions of grammaticalization and their problems. *Language Sciences*. 23. 93-112.
- Christofalos, Andriana & Feldman, Laurie & Sheridan, Heather. (2022). Semantic Congruency Facilitates Memory for Emojis.
- Cohn, N., Roijackers, T., Schaap, R., & Engelen, J. (2018). Are emoji a poor substitute for words? Sentence processing with emoji substitutions. In *Proceedings of the 40th Annual Conference of the Cognitive Science Society* (pp. 1524-1529). The Cognitive Science Society.
- Cohn, Neil & Engelen, Jan & Schilperoord, Joost. (2019). The grammar of emoji? Constraints on communicative pictorial sequencing. *Cognitive Research: Principles and Implications*. 4.
- Collins, Chris. (2007). *Home Sweet Home*. Manuscript. NYU.
- Collins, C. & Kayne, R. (2021) *Towards a Theory of Morphology as Syntax*. Manuscript. Lingbuzz
- Curtis, Charles. (2017). Just when did we all start using GOAT anyway? *For The Win*. USA Today. Blog
- Dahl, E., & Fábregas, A. (2018). Zero Morphemes. *Oxford Research Encyclopedia of Linguistics*.
- Danesi, M. (2017). *The semiotics of emoji*. London: Bloomsbury Academic
- De Belder, Marijke & Faust, Noam & Lampitelli, Nicola. (2014). On a low and a high diminutive.
- Embick, David & Marantz, Alec. (2008). Architecture and Blocking. *Linguistic Inquiry - LINGUIST INQ*. 39.
- Esipova, Maria. (2019). Composition and projection of co-speech gestures. *Semantics and Linguistic Theory*. 29. 117.
- Evans, V. (2017). *The Emoji code: The linguistics behind smiley faces and scaredy cats*.
- Fahlman, Scott. (1982). *Smiley Lore :-)*. Carnegie Melon University.
- Gawne, L., & McCulloch, G. (2019). *Emoji as digital gestures*.

- Giegerich, H. J. (1999). *Lexical strata in English: Morphological causes, phonological effects*. Cambridge Studies in Linguistics 89. Cambridge, UK: Cambridge University Press.
- Giegerich, H. J. (2005). Lexicalism and modular overlap in English. *SKASE Journal of Theoretical Linguistics*. 2(2). 43-62.
- Greenberg, G. (2021). Semantics of Pictorial Space. *Review of Philosophy and Psychology*, 1-41.
- Grosz, P.G., Greenberg, G., Leon, C.P., & Kaiser, E. (2022). A semantics of face emoji in discourse.
- Grosz, P.G., Kaiser, E., & Pierini, F. (2021). Discourse anaphoricity and first-person indexicality in emoji resolution.
- Hilpert, M. (2020). Lexicalization in Morphology. In *Oxford Research Encyclopedia of Linguistics* (pp. 1-18). Oxford: Oxford University Press
- Hunter, J. (2019). Relating Gesture to Speech: reflections on the role of conditional presuppositions. *Linguist and Philos.* 42. 317–332
- Ishmael, Aiyana. (2021). Sending Smiley Emojis? They Now Mean Different Things to Different People. Wall Street Journal.
- Kastovsky, Dieter. (1994). Typological differences between English and German morphology and their causes. In T. Swan, E. Mørck & O. Westvik (Ed.), *Language Change and Language Structure: Older Germanic Languages in a Comparative Perspective* (pp. 135-158). Berlin, New York: De Gruyter Mouton.
- Kiparsky, P. (2015). Stratal OT: A synopsis and FAQs.
- Kiparsky, P. (2020). Morphological Units: Stems. In *Oxford Research Encyclopedia of Linguistics*. Oxford: Oxford University Press
- Knudson, A. (2017). Health department STD campaign uses emoji lingo. Staten Island Live
- Kramer, R. (2016) The location of gender features in the syntax. *Language and Linguistics Compass*, 10: 661– 677.
- Labov, W. (1966). *The Social Stratification of English in New York City*. Cambridge University Press.
- Ladislades, J. (1635). Financial Record.

- Lasnik, H. 1995. Verbal morphology: Syntactic Structures meets the minimalist program. In H. Campos and P. Kempchinsky (eds.) *Evolution and Revolution in Linguistic Theory: Essays in Honor of Carlos Otero*. Georgetown University Press, pp. 251-275.
- Lauwers, Peter & Willems, Dominique. (2011). Coercion: Definition and challenges, current approaches, and new trends. *Linguistics*. 49.
- Leondis, T. (2017). *The Emoji Movie*.
- Lieber, Rochelle. (1981). *On the organization of the lexicon*. Bloomington: Indiana University Linguistic Club
- Londhal, T. (2020). Syntactic Categorization of Roots. In *Oxford Research Encyclopedia of Linguistics*. Oxford: Oxford University Press
- Maier, E. (2020). Emojis as Pictures
- Marantz, A. (1997). No escape from syntax: Don't try morphological analysis in the privacy of your own lexicon.
- Marantz, Alec. (2013). Locality Domains for Contextual Allomorphy across the Interfaces. 95-116.
- Michaelis, Laura. (2004). Type shifting in construction grammar: An integrated approach to aspectual coercion. *Cognitive Linguistics*.
- Migott, L. & Guerrini, J. (2022). Linguistic Inferences from Pro-Speech Music. *Lingbuzz*.
- Moschini, I. (2016). The "Face with Tears of Joy" Emoji. A Socio-Semiotic and Multimodal Insight into a Japan-America Mash-Up. *HERMES - Journal of Language and Communication in Business*, 55, 11-25.
- Nida, E. (1948). The identification of morphemes. *Language*, 24, 414-441.
- Olsen, Susan. (1986). *Wortbildung im Deutschen. Eine Einführung in die Theorie der Wortstruktur*. Stuttgart: Kröner.
- Pasternak, R., & Tieu, L. (2022). Co-linguistic content inferences: From gestures to sound effects and emoji. *Quarterly Journal of Experimental Psychology*.
- Picallo, M.C. 1991. Nominals and nominalization in Catalan. *Probus* 3. 279-316.
- Pierini, F. (2021). Emojis and gestures: a new typology. *Proceedings of Sinn Und Bedeutung*, 25, 720-732.

- Roberts, Julie. (2003). Style and Sociolinguistic Variation. *American Anthropologist*. 105. 10.1525/aa.2003.105.4.856.1.
- Roberts, I., & Roussou, A. (1999). *Syntactic Change: A Minimalist Approach to Grammaticalization*.
- Schlenker, Philippe. (2018). Iconic Pragmatics. *Natural Language & Linguistic Theory*. 36.
- Schlenker, P. (2018). Gestural Semantics. *Nat Lang Linguist Theory*. 37. 735-784.
- Schlenker, Philippe. (2019). What is Super Semantics?*. *Philosophical Perspectives*. 32.
- Schoenmakers, Gert-Jan & Storment, John David. (2021). Going city: Directional predicates and preposition incorporation in youth vernaculars of Dutch. *Linguistics in the Netherlands*. 38. 65-80.
- Selkirk, Elisabeth. (1982). *The syntax of words*. Cambridge, Mass.: M.I.T. Press.
- Siegel, D. (1974). *Topics in English morphology* (Doctoral thesis). MIT, Cambridge, MA. Published in 1979 by Garland Press, New York, NY.
- Stahl, L. (2014). The First Emoticon? *Ivebeenreadinglately*. Blog.
- Sweetser, E. (1988). *Grammaticalization and Semantic Bleaching*.
- Trudgill, P. (1972). Sex, covert prestige and linguistic change in the urban British English of Norwich. *Language in Society*, 1(2), 179-195.
- Votruba, M. (2018). 17th Century Emoji. Slovak Studies Program. University of Pittsburgh.
- Weissman, Benjamin. (2019). Emojis in Sentence Processing: An Electrophysiological Approach. WWW '19: Companion Proceedings of The 2019 World Wide Web Conference. 478-479.
- Zwicky, A. M. (1992). Some choices in the theory of morphology. In *Formal grammar: Theory and implementation* (p.p. 327-371). Vancouver Studies in Cognitive Science 2. Oxford, UK Oxford University Press.