A formal linguistic approach to affective emojis in CMC

Chenchen (Julio) Song

Zhejiang University

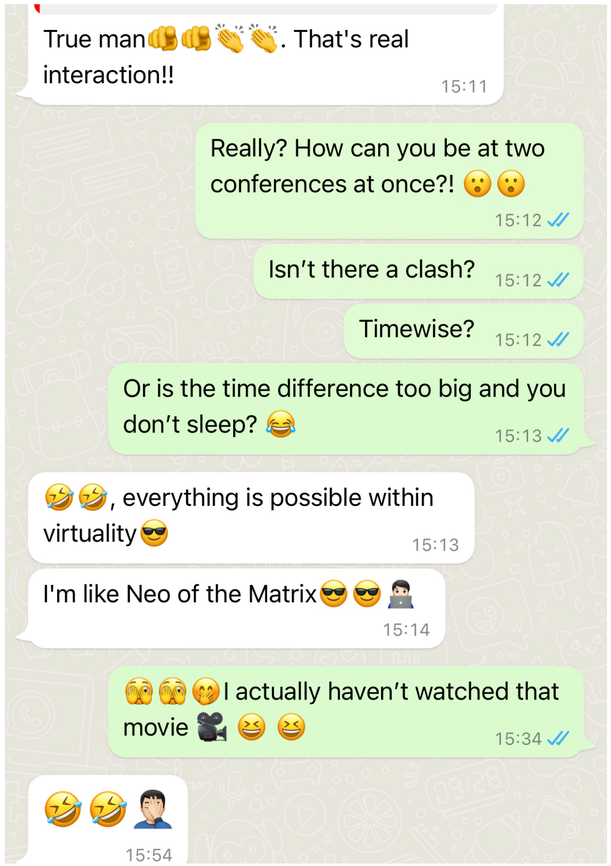
The 2022 Seoul International Conference on Linguistics (SICOL-2022)
August 11–12, 2022

Emojis in CMC (=computer-mediated communication)

An increasingly important part of life

A recent WhatsApp chat of mine 👇

(interlocutor consent obtained)



"92% of the world's online population use emoji"

— **Jennifer Daniel**, **Unicode Emoji Subcommittee Chair**

Two main uses of emojis

Affective vs. nonaffective

👆aka “non-at-issue” 👆aka “at-issue”

(Potts’s 2005 et seq. terminology)

Example:

(1) a. Great idea 👍 I’m in 😊 affective

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

I focus on the affective use in this talk.

Affective emojis

Characteristics

1. Conveying tones or emotions
2. Mainly face emojis, but also many nonface ones
3. Typically appended to sentences at the end

Affective emojis

Characteristics

1. Conveying tones or emotions Great idea 👍 I’m in 😊
2. Mainly face emojis, but also many nonface ones
3. Typically appended to sentences at the end

Affective emojis

Characteristics

1. Conveying tones or emotions Great idea 👍 I’m in 😊
2. Mainly face emojis, but also many nonface ones
3. Typically appended to sentences at the end

Top 10 emojis used worldwide: 😂 ❤️ 🤣 👍 😭 🙏 😘 🥰 😍 😊

— **Unicode Consortium (2021)**

Affective emojis

Characteristics

1. Conveying tones or emotions Great idea 👍 I’m in 😊
2. Mainly face emojis, but also many nonface ones
3. Typically appended to sentences at the end

Top 10 emojis used worldwide: 😂 ❤️ 🤣 👍 😭 🙏 😘 🥰 😍 😊

— **Unicode Consortium (2021)**

Example: 💅 is often used to display an air of nonchalance or indifference (Emojipedia)
(2) As i said before, u can't compare urself with us. We're on another level, we're on the next level. Sorry to say, but it's a fact 💅 (Twitter)

Questions

Affective emojis' place in CMC grammar

1. Do they have a generative syntax?
2. Do they have a model-theoretic semantics?
3. What does research on CMC grammar entail? (big picture)

Questions

Affective emojis' place in CMC grammar

1. Do they have a generative syntax?
2. Do they have a model-theoretic semantics?
3. What does research on CMC grammar entail? (big picture)

Main proposal: Affective emojis are a (semi)lexical category in CMC.

Questions

Affective emojis' place in CMC grammar

1. Do they have a generative syntax?
2. Do they have a model-theoretic semantics?
3. What does research on CMC grammar entail? (big picture)

Main proposal: Affective emojis are a (semi)lexical category in CMC.

Syntax → **Semantics**

Generalized Root Syntax (Song 2019)

Monadic Composition (Song 2021b)

Road map

1. Affective emojis as a (semi)lexical category
2. Syntax (& big picture issues)
3. Semantics

1. Affective emojis are a (semi)lexical category

Lexical status: an open class

1. New face emojis are created every year

2018 

2019 

2020 

2021/22 

What next?

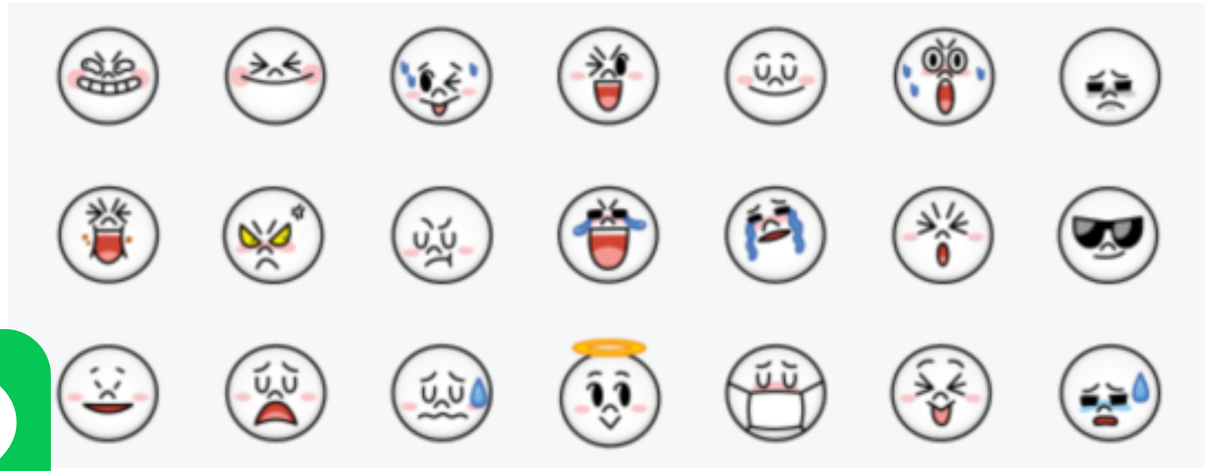
Lexical status: an open class

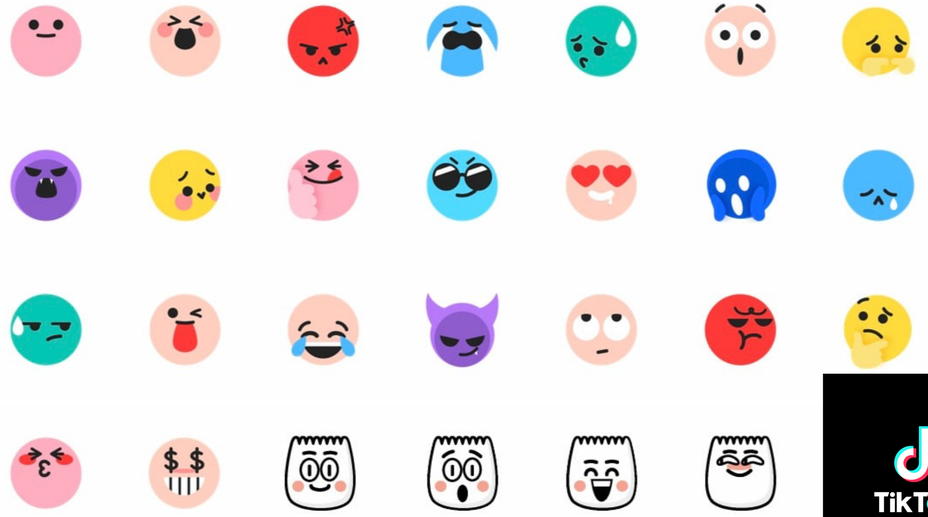
2. Many platform-specific ones

Weibo

WeChat

(see emojiall.com for more)





Lexical status: an open class

3. Many nonface emojis can be used affectively too

Lexical status: an open class

3. Many nonface emojis can be used affectively too

👏🤝👍💅🙏☕🚬🎉💣❤️💝🌹…

4. Various quasi emojis

emoticons/kaomojis

Highly popular and versatile in Asia

https://pin.it/pZZtx4t

Grammatical status: a clearly fixed function

Add tones or speaker attitudes to linguistic content

Similar to the function of certain particles in non-CMC linguistics

- Chinese sentence-final particles (SFPs)
- German sentence-middle (aka modal) particles

Grammatical status: a clearly fixed function

Add tones or speaker attitudes to linguistic content

Similar to the function of certain particles in non-CMC linguistics

- Chinese sentence-final particles (SFPs)
- German sentence-middle (aka modal) particles

(3) a. *xià xuě le ye* [Mandarin Chinese]
fall snow PRF SFP
"It's snowing. (happy tone)"

b. *xià xuě le a*
fall snow PRF SFP
"It's snowing. (surprised tone)"

c. *xià xuě le you*
fall snow PRF SFP
"It's snowing. (kind reminder tone)"

Grammatical status: a clearly fixed function

Add tones or speaker attitudes to linguistic content

Similar to the function of certain particles in non-CMC linguistics

- Chinese sentence-final particles (SFPs)
- German sentence-middle (aka modal) particles

German modal particles

- used mainly in the spontaneous spoken language in colloquial registers in German
- reflect the mood or the attitude of the speaker
- highlight the sentence's focus

Example	Connotation
halt, nun, einmal	some unpleasant fact must be accepted
ja	reminder to the listener
mal	a casual, less blunt tone
doch	emphasis, urgency, impatience, etc. (highly versatile)

(Wikipedia)

Grammatical status: a clearly fixed function

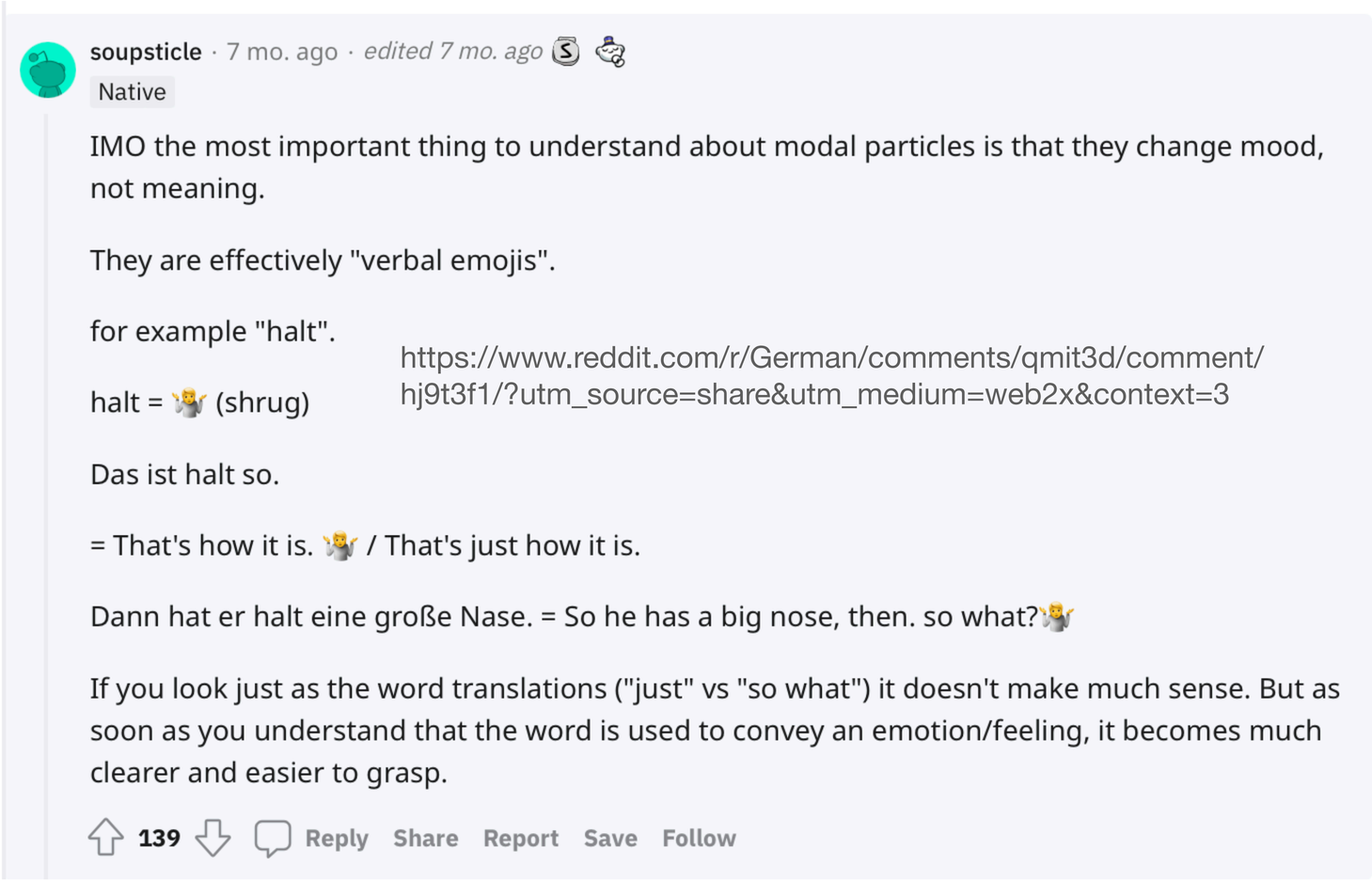
Add tones or speaker attitudes to linguistic content

Similar t … **stics**

- Chines
- Germa

German … **rman**

- used
- reflec
- highl



soupsticle · 7 mo. ago · *edited 7 mo. ago*

Native

IMO the most important thing to understand about modal particles is that they change mood, not meaning.

They are effectively "verbal emojis".

for example "halt".

halt = 🤷 (shrug)

https://www.reddit.com/r/German/comments/qmit3d/comment/hj9t3f1/?utm_source=share&utm_medium=web2x&context=3

Das ist halt so.

= That's how it is. 🤷 / That's just how it is.

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

If you look just as the word translations ("just" vs "so what") it doesn't make much sense. But as soon as you understand that the word is used to convey an emotion/feeling, it becomes much clearer and easier to grasp.

139 Reply Share Report Save Follow

…satile)

(Wikipedia)

Grammatical status: a clearly fixed function

Add tones or speaker attitudes to linguistic content

Similar to the function of certain particles in non-CMC linguistics

- Chinese sentence-final particles (SFPs)
- German sentence-middle (aka modal) particles

German modal particles

- used mainly in the spontaneous spoken language in colloquial registers in German
- reflect the mood or the attitude of the speaker
- highlight the sentence's focus

Example	Connotation
halt, nun, einmal	some unpleasant fact must be accepted
ja	reminder to the listener
mal	a casual, less blunt tone
doch	emphasis, urgency, impatience, etc. (highly versatile)

(Wikipedia)

Lexical status + grammatical status

= a semilexical (aka semifunctional) category

Some typical semilexical categories (see Song 2021a for an overview):

	Grammatical function	Lexical idiosyncrasy	Example
Classifiers	Atomizing mass concepts and making them countable	Various conventionalized perspectives	(Mandarin) *zhī* for long, think objects, *bǎ* for objects with handle-like parts, etc.
Light verbs	Various event structure functions	Lexical selection, register variation, etc.	(English) *take* a shower, *do* the laundry, *make* a phone call, etc.
Adpositions	Additional predication	Various concrete (e.g., spatial) relations	(English) *in*, *on*, *at*, *below*, etc.

Affective emojis are semilexical

	Grammatical function	Lexical idiosyncrasy	Example
	Add tones/attitudes to linguistic content	happy, loving	I’m in 😊
		praising, supportive	Great idea 👍
		nonchalant, arrogant	Sorry to say, but it's a fact 💅
		gossipy (specific to Chinese)	‘Just found out that Wahaha had changed their endorser from Leehom Wang to Greg Han.🥵’

Affective emojis are semilexical

	Gran
😊	
👍	Add
💅	lir
🥲	

Don't put on a happy face! Are you using the smiley emoji all wrong?

The classic grinning emoji has once more changed its meaning – at least amongst gen Zers. So what is it communicating now – and what should you be using instead?

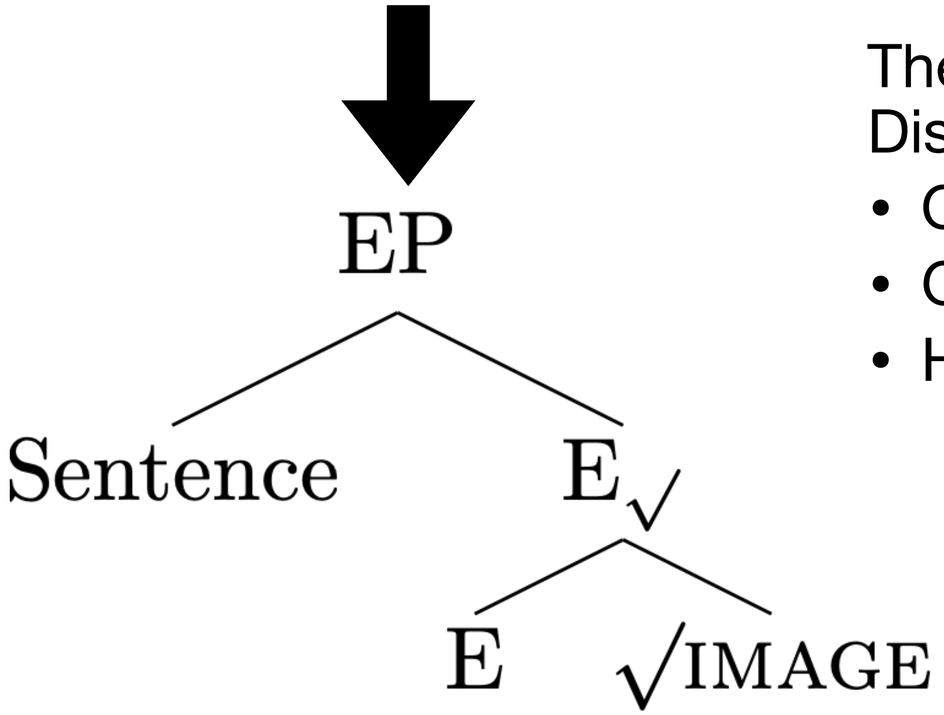
The Guardian

Great idea 👍
rry to say, but it's a fact 💅
out that Wahaha had changed their om Leehom Wang to Greg Han.🥲'

2. Syntax

The emotional shell category E

[EP Sentence [E E √IMAGE]] (an updated version of Song 2019)

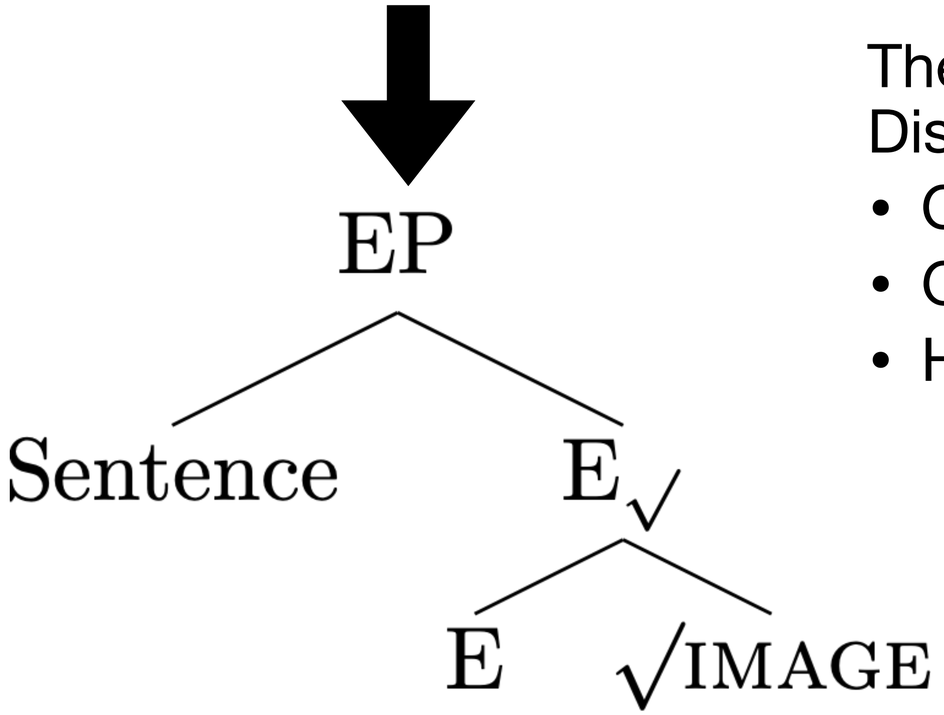


The root categorization technique is borrowed from Distributed Morphology (Halle & Marantz 1993 et seq.):

- Originally used for content word formation
- Generalized to semilexical words in Song 2019
- Here modeling the lexical side of affective emojis

The emotional shell category E

[EP Sentence [E E √IMAGE]] (an updated version of Song 2019)



The root categorization technique is borrowed from Distributed Morphology (Halle & Marantz 1993 et seq.):

- Originally used for content word formation
- Generalized to semilexical words in Song 2019
- Here modeling the lexical side of affective emojis

The grammatical category E functions like an emotional wrapper for the linguistic sentence.

Each affective emoji is a tiny “idiom” in the CMC lexicon.

The emotional shell category E

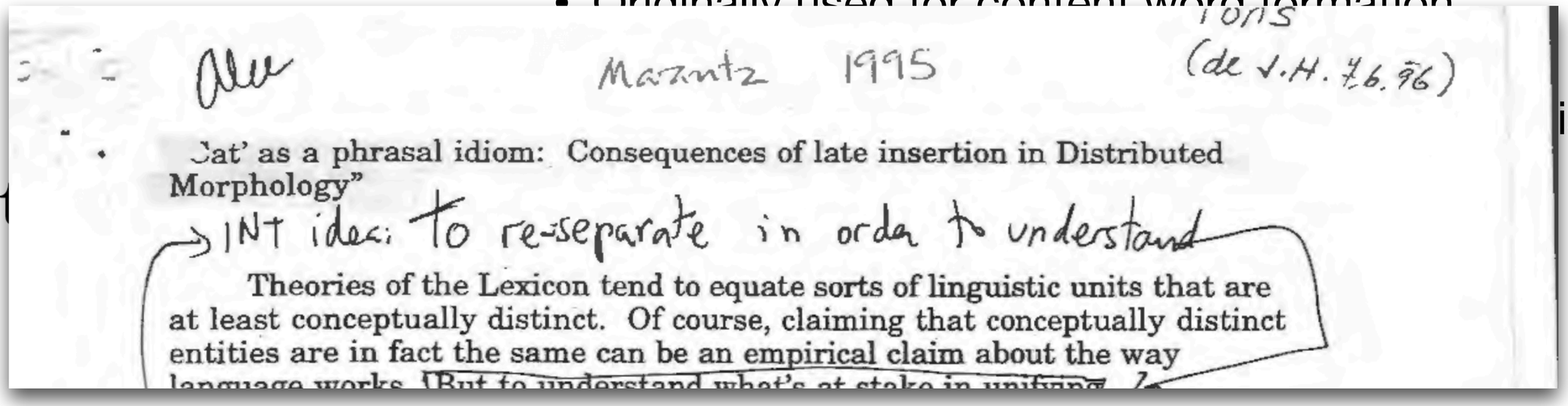
[EP Sentence [E E √IMAGE]] (an updated version of Song 2019)

The root categorization technique is borrowed from Distributed Morphology (Halle & Marantz 1993 et seq.):

- Originally used for content word formation

is

Sent

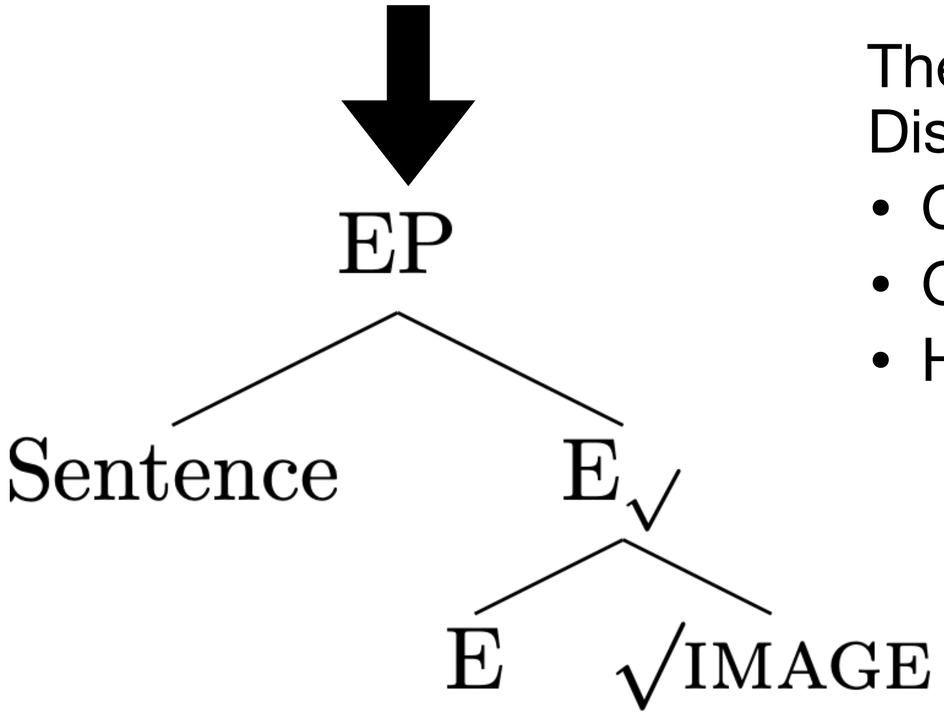


The grammatical category E functions like an emotional wrapper for the linguistic sentence.

Each affective emoji is a tiny "idiom" in the CMC lexicon.

The emotional shell category E

[EP Sentence [E E √IMAGE]] (an updated version of Song 2019)



The root categorization technique is borrowed from Distributed Morphology (Halle & Marantz 1993 et seq.):

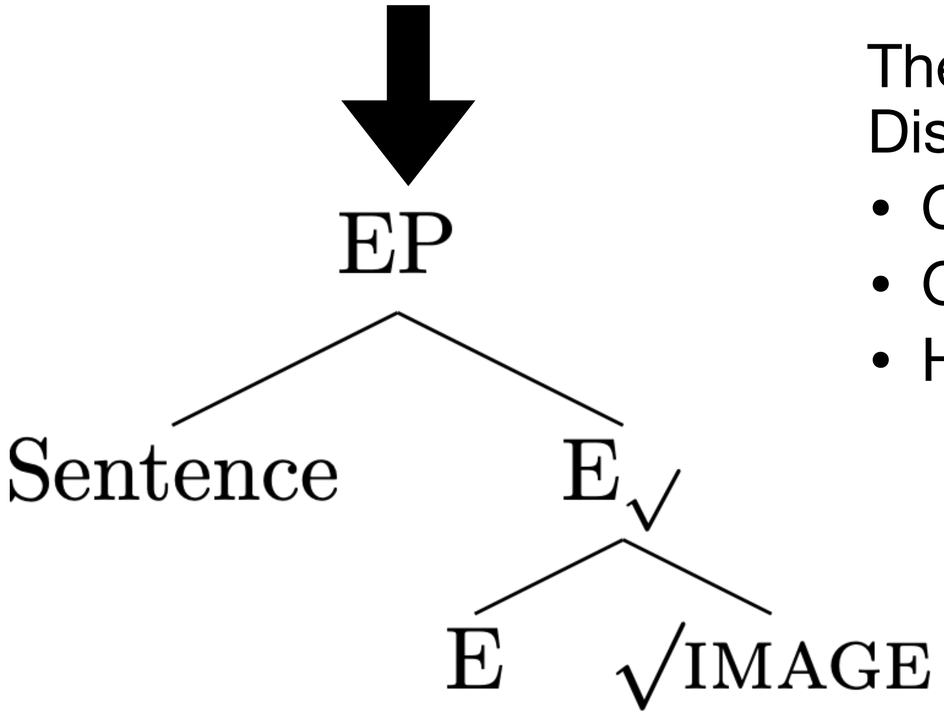
- Originally used for content word formation
- Generalized to semilexical words in Song 2019
- Here modeling the lexical side of affective emojis

The grammatical category E functions like an emotional wrapper for the linguistic sentence.

Each affective emoji is a tiny “idiom” in the CMC lexicon.

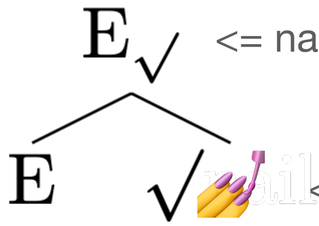
The emotional shell category E

[EP Sentence [E E √IMAGE]] (an updated version of Song 2019)



The root categorization technique is borrowed from Distributed Morphology (Halle & Marantz 1993 et seq.):

- Originally used for content word formation
- Generalized to semilexical words in Song 2019
- Here modeling the lexical side of affective emojis



<= nail polish image used affectively (idiomatic: nonchalance)

<= nail polish image

The grammatical category E functions like an emotional wrapper for the linguistic sentence.

Each affective emoji is a tiny “idiom” in the CMC lexicon.

Affective emojis vs. affective particles

Why can’t we model them with a single functional category?

Two main reasons:

1. They can and often do co-occur.
2. The positioning of affective emojis is not affected by crosslinguistic word order variation, while that of affective particles is.

(examples from Weibo, the Chinese equivalent of Twitter)

Positioning of affective emojis

A survey of nine languages on social media websites (Twitter, Weibo)

Language	Family	Type	Basic word order	Place of affective emoji
Mandarin	Sinitic	isolating	SVO	sentence-final
Japanese	Japonic	agglutinative	SOV	sentence-final
Korean	Koreanic	agglutinative	SOV	sentence-final
English	Germanic	analytic	SVO	sentence-final
German	Germanic	fusional	SOV (V2 in matrix)	sentence-final
French	Romance	fusional	SVO	sentence-final
Irish	Celtic	fusional	VSO	sentence-final
Basque	Language isolate	agglutinative/ fusional	SOV	sentence-final
Hungarian	Finno-Ugric	agglutinative	relatively free	sentence-final

Positioning of affective emojis

A survey of nine languages on social media websites (Twitter, Weibo)

Example: (from Twitter)

(5) a. *Les pères, ils ont droit au whisky et autres alcool de “bonhomme”* 🤦‍♀️ [French]
“The fathers, they have the right to whisky and other alcohols of ‘fellow.’”

b. *Ich dachte immer, dass hier alles anonym ist* 🤷‍♂️😉 [German]
“I always thought that everything was anonymous here.”

c. *gozenchū no ame wa dokoni ittandesu ka* 🤔 [Japanese]
“Where did the rain in the morning go?”

d. *Membeo-deul-i ‘hat-gyu’-rago bureum* 🥹 [Korean]
“The members calling him ‘hot-gyu’”

e. *RT agus fág trácht le bheith san áireamh!!* 😍 [Irish]
“RT and leave a comment to be included!!”

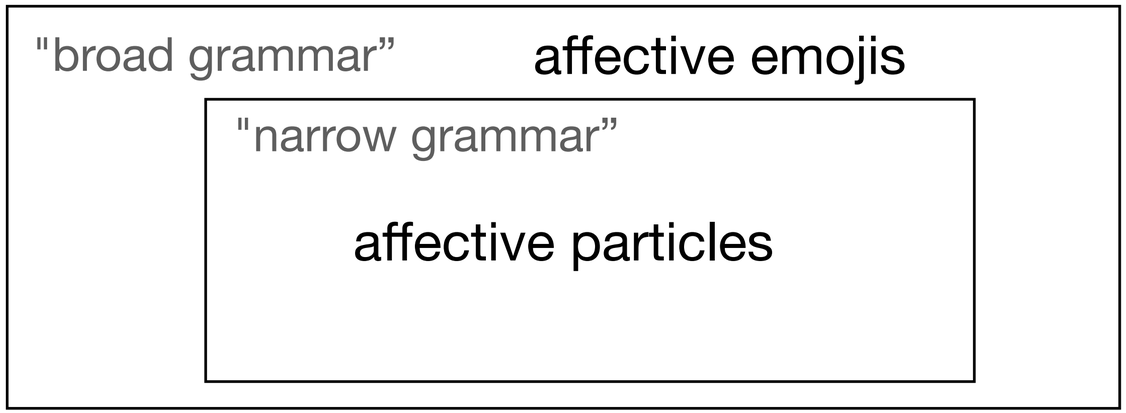
f. *Bilera eta ekitaldi nagusiak bueltan dira Euskaldunan* 😊 [Basque]
“Meetings and big events are back in Basque.”

g. *Legyetek a barátaim, ugyanígy doraszell a nevem* 😊 [Hungarian]
“Be my friends (on BeReal). My name is just doraszell.”

Interim summary

Affective emojis and affective particles are not introduced by the same functional category

Intuitively, affective particles are still **within** the linguistic content, whereas affective emojis are **outside** of it, functioning like a **higher-order shell** specific to CMC grammar.



Interim summary

Not just emojis, but various other visual elements can be used affectively too!

is for me?

These are part of the “broad grammar” of CMC but not part of the “narrow grammar” of traditional linguistics.

Interim summary

Not just emojis, but various other visual elements can be used affectively too!

is for me?

These are part of the “broad grammar” of CMC but not part of the “narrow grammar” of traditional linguistics.

Interim summary

On some platforms, even sound elements may be used affectively

A recent Instagram story of mine 👇

Here, the upbeat/enthusiastic tone accompanying the post is conveyed through the background music — similar in effect to 👇

Yummy cherry tomatoes for HOT summer days! 😎🐓🏝️

Interim summary

On some platforms, even sound elements may be used affectively

A recent Instagram story of mine 👇

Here, the upbeat/enthusiastic tone accompanying the post is conveyed through the background music — similar in effect to 👇

Yummy cherry tomatoes for HOT summer days! 😎🐓🏝️

Interim summary

On some platforms, even sound elements may be used affectively

A recent Instagram story of mine 👇

Here, the upbeat/enthusiastic tone accompanying the post is conveyed through the background music — similar in effect to 👇

Yummy cherry tomatoes for HOT summer days! 😎🐓🏝️

=> The emotional wrapper in the "broad grammar" of CMC can recycle miscellaneous multimedia elements affectively.

From “narrow” to “broad” grammar

What formal linguistic tools are safe to use?

- We want to apply formal linguistic tools to not-entirely-linguistic data.
- This is fine as long as the tools are sufficiently **domain-general**.
- Basically, anything motivated by “interface conditions” in current generative syntax (e.g., Move, Agree, Phases) risks being domain-specific.

From “narrow” to “broad” grammar

What formal linguistic tools are safe to use?

- We want to apply formal linguistic tools to not-entirely-linguistic data.
- This is fine as long as the tools are sufficiently **domain-general**.
- Basically, anything motivated by “interface conditions” in current generative syntax (e.g., Move, Agree, Phases) risks being domain-specific.

Some “safe” tools

✓Merge basic combinatorial operation

✓Categorization recycling existing material for new purpose

✓Model-theoretic semantics not limited to natural languages

From “narrow” to “broad” grammar

What formal linguistic tools are safe to use?

- We want to apply formal linguistic tools to not-entirely-linguistic data.
- This is fine as long as the tools are sufficiently **domain-general**.
- Basically, anything motivated by “interface conditions” in current generative syntax (e.g., Move, Agree, Phases) risks being domain-specific.

Some “safe” tools

✓Merge basic combinatorial operation

✓Categorization recycling existing material for new purpose

✓Model-theoretic semantics not limited to natural languages

Bottom line: CMC data are amenable to symbolic analysis.

3. Semantics

Formal semantics for (Generalized) Root Syntax

via monadic composition (Song 2021b, 2022a)

Recall:

- Root Syntax keeps purely functional and idiosyncratic information apart.
- The root categorization schema holds the two types of info together.

Formal semantics for (Generalized) Root Syntax

via monadic composition (Song 2021b, 2022a)

Recall:

- Root Syntax keeps purely functional and idiosyncratic information apart.
- The root categorization schema holds the two types of info together.

Desideratum:

- Semantic composition should mirror the above syntactic mechanism.

Formal semantics for (Generalized) Root Syntax

via monadic composition (Song 2021b, 2022a)

Recall:

- Root Syntax keeps purely functional and idiosyncratic information apart.
- The root categorization schema holds the two types of info together.

Desideratum:

- Semantic composition should mirror the above syntactic mechanism.

The writer monad (via Asudeh & Giorgolo 2020, originally from CS/Math):

- Maps each pure-function denotation ⟦F⟧ to an "enriched" type ⟨⟦F⟧, {…}⟩.
- The enriching mechanism relies on established properties of the universe of sets.
- The monad systematically keeps pure-function composition and idiosyncratic enrichment apart via the ordered pair structure.

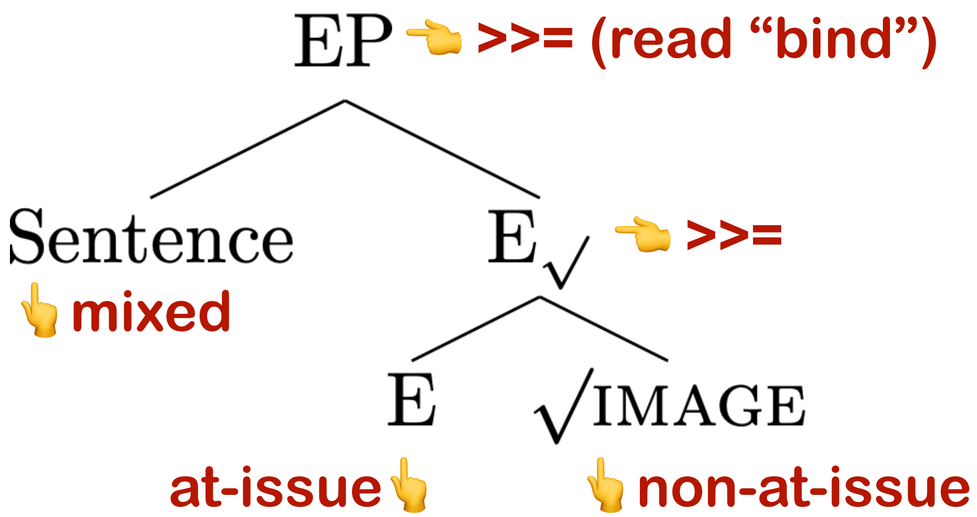
Some background

Modes of composition

1. Function application => most often used
   input: f, x; output: f(x)
2. Conjunction => used for “predicate modification” and event semantics
   input: f, g; output: f&g
3. Monadic “bind” => used for “nonpure” computations with “side effects”
   input: f*, x*; output: f(x)** [I use the superscript * to indicate side effect]
   pure computation: f(x); nonpure/side effects: **

1 and 2 are already available in Heim & Kratzer (1998)
3 originates in mathematical category theory and functional programming but has been introduced to linguists too (Shan 2002; Asudeh & Giorgolo 2020; Song 2021b, 2022a)

Root Syntax 🤝 Monadic Semantics



Template:

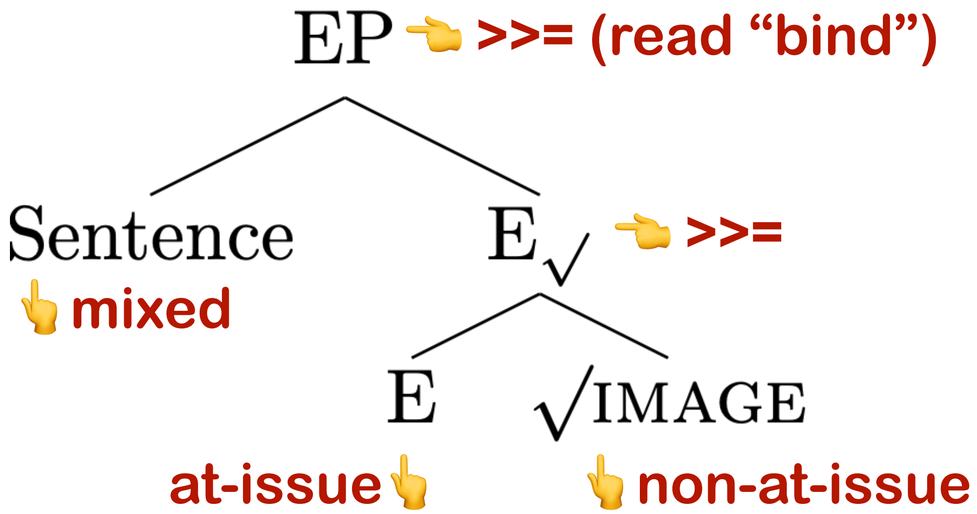
$[_{>>=}$ X YP]

= write(⟦YP⟧) >>= λy.η(⟦X⟧)

= ⟨⟦X⟧(⟦YP⟧), $NAI_X \cup NAI_{YP}$}⟩

(NAI = non-at-issue content)

Root Syntax 🤝 Monadic Semantics



1. ⟦E⟧ = λxλu . {w | AFFECT(x, u) at w}

(adapted from Grosz et al. 2021)
(see Song 2022a for more detail)

x affectively performs the speech act of u at w

Template:

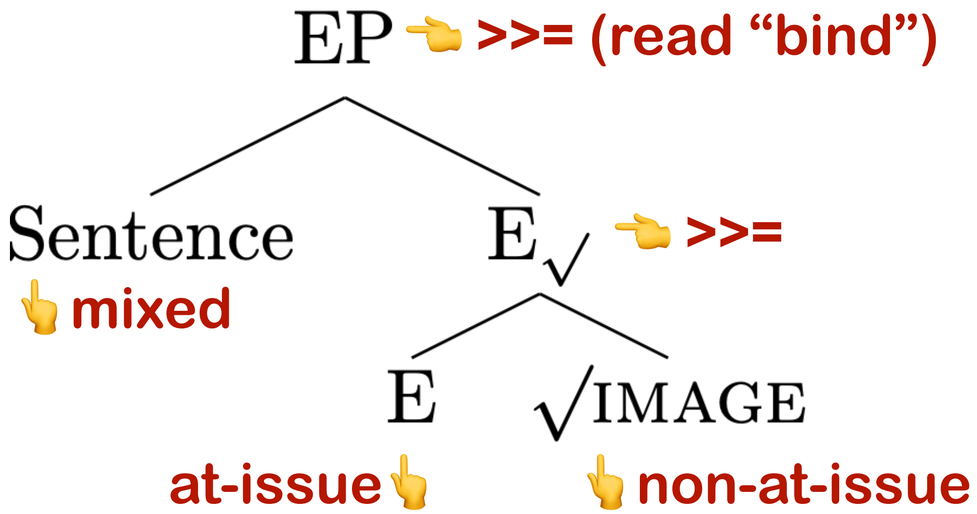
$[_{>>=}$ X YP]

= write(⟦YP⟧) >>= λy.η(⟦X⟧)

= ⟨⟦X⟧(⟦YP⟧), $NAI_X \cup NAI_{YP}$⟩

(NAI = non-at-issue content)

Root Syntax 🤝 Monadic Semantics



1. ⟦E⟧ = λxλu . {w | AFFECT(x, u) at w}

(adapted from Grosz et al. 2021) 👆

(see Song 2022a for more detail)

x affectively performs the speech act of u at w

2. ⟦E√⟧ = ⟦[E E √IMAGE]⟧

= write(E√) >>= λy.η(⟦E⟧)

= ⟨⟦E⟧, {E is enriched by √IMAGE}⟩

Template:

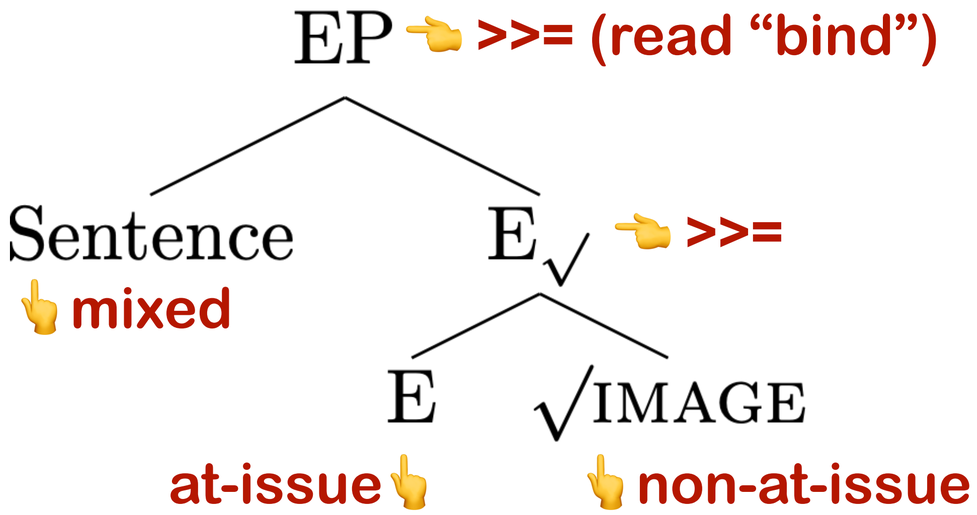
[$_{>>=}$ X YP]

= write(⟦YP⟧) >>= λy.η(⟦X⟧)

= ⟨⟦X⟧(⟦YP⟧), $NAI_X \cup NAI_{YP}$}⟩

(NAI = non-at-issue content)

Root Syntax 🤝 Monadic Semantics



1. ⟦E⟧ = λxλu . {w | AFFECT(x, u) at w}
(adapted from Grosz et al. 2021) 👆
(see Song 2022a for more detail)

x affectively performs the speech act of u at w

2. ⟦E√⟧ = ⟦[E E √IMAGE]⟧
= write(E√) >>= λy.η(⟦E⟧)
= ⟨⟦E⟧, {E is enriched by √IMAGE}⟩

3. ⟦EP⟧ = write(⟦E√⟧) >>= λy.η(⟦Sentence⟧)
= write(⟦[E E √IMAGE]⟧) >>= λy.η(⟦Sentence⟧)
= write(⟨⟦E⟧, {E is enriched by √IMAGE}⟩) >>= λy.η(⟦Sentence⟧)
= ⟨⟦E⟧(⟦Sentence⟧), {…E is enriched by √IMAGE…}⟩

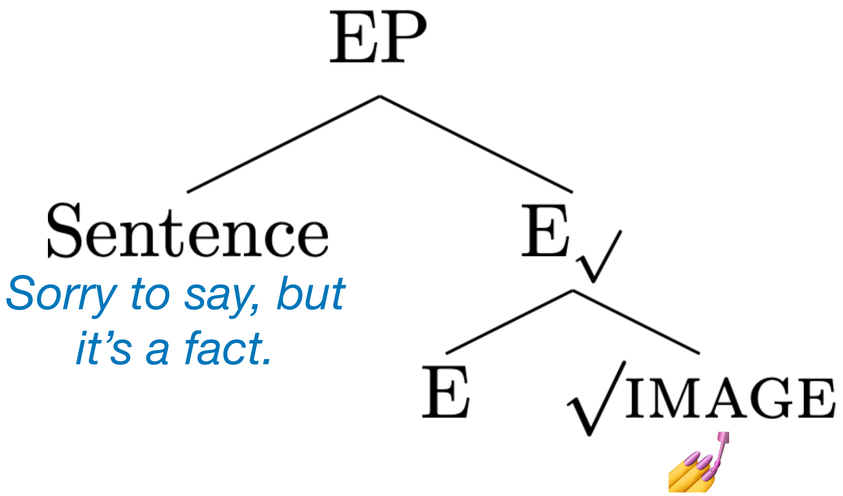
👆pure-function composition

👆accumulated idiosyncratic enrichment

Template:
[>>= X YP]
= write(⟦YP⟧) >>= λy.η(⟦X⟧)
= ⟨⟦X⟧(⟦YP⟧), $NAI_X \cup NAI_{YP}$⟩
(NAI = non-at-issue content)

Illustration

Example 1: Sorry to say, but it's a fact. 💅



The speaker (S) performs a declarative speech act in a tone conventionalized by the affective recycling of this image: 💅.

⟦EP⟧ = write(⟦E√⟧) >>= λy.η(⟦Sorry to say, but it's a fact.⟧)

= …

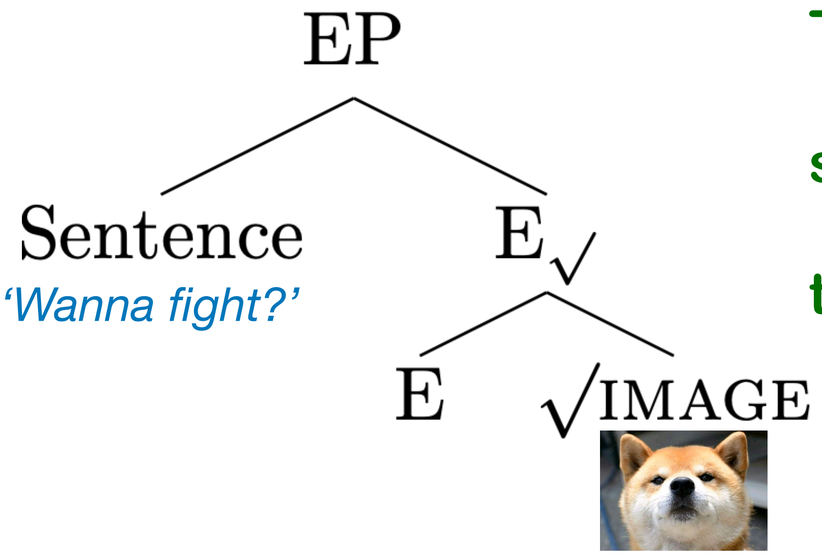
=⟨(λu . {w | AFFECT(S, u) at w})(⟦Sorry….⟧), {…E is enriched by 💅…}⟩

=⟨{w | AFFECT(S, ⟦Sorry….⟧) at w}, {…**nonchalant tone**…}⟩

👆**at-issue** 👆**non-at-issue**

Illustration

Example 2

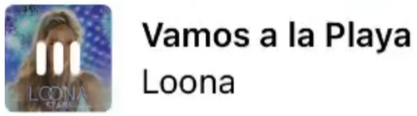


The speaker (S) performs an interrogative speech act in a tone conventionalized by the affective recycling of this image: 

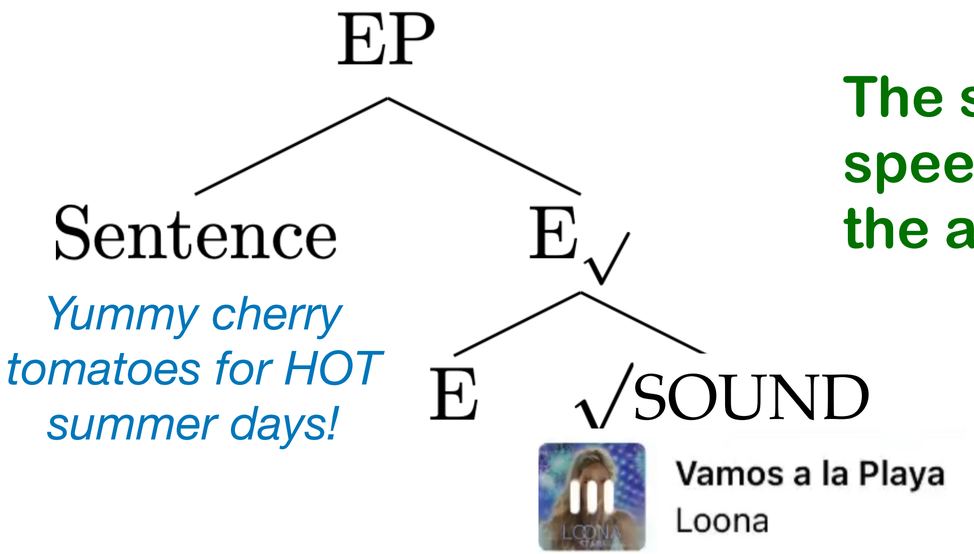
⟦EP⟧ = write(⟦E√⟧) >>= λy.η(⟦‘Wanna fight?’⟧)
= …
=⟨(λu . {w | AFFECT(S, u) at w})(⟦‘Wanna fight?’⟧), {…E is enriched by  …}⟩
=⟨{w | AFFECT(S, ⟦‘Wanna fight?’⟧) at w}, {…**jocularly menacing tone**…}⟩

👆 **at-issue** 👆 **non-at-issue**

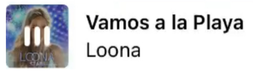
Illustration



Example 3: Yummy cherry tomatoes for HOT summer days!



The speaker (S) performs a declarative speech act in a tone conventionalized by the affective recycling of this sound file:



⟦EP⟧ = write(⟦E√⟧) >>= λy.η(⟦Yummy…⟧)

= …

=⟨(λu . {w | AFFECT(S, u) at w})(⟦Yummy…⟧), {…E is enriched by  …}⟩

=⟨{w | AFFECT(S, ⟦Yummy…⟧) at w}, {…**upbeat tone**…}⟩

👆**at-issue** 👆**non-at-issue**

Conclusion

Questions (repeated)

Affective emojis' place in CMC grammar

1. Do they have a generative syntax?
2. Do they have a model-theoretic semantics?
3. What does research on CMC grammar entail? (big picture)

Main proposal: Affective emojis are a (semi)lexical category in CMC.

Syntax	→	**Semantics**
Generalized Root Syntax (Song 2019)		Monadic Composition (Asudeh & Giorgolo 2020, Song 2021b)

Results

A formal linguistic approach to affective emojis in CMC

1. Affective emojis are a semilexical category in CMC.
2. Their syntax can be modeled by Generalized Root Syntax.
3. Their formal semantics can be modeled by the writer monad.
4. We can apply the domain-general subset of formal linguistic tools to the "broad grammar" of CMC.

CMC forces us to think outside of the conventional linguistics box!

In this talk, I used

- ✓Merge basic combinatorial operation
- ✓Categorization recycling existing material for new purpose
- ✓Model-theoretic semantics not limited to natural languages

Thank you!

References

Asudeh, A. & G. Giorgolo (2020). *Enriched meanings*. OUP.

Grosz, P., G. Greenberg, C. De Leon & E. Kaiser (2021). A semantics of face emoji in discourse. Ms.

Halle, M. & A. Marantz (1993). Distributed morphology and the pieces of inflection. In *Essays in linguistics in honor of Sylvain Bromberger*, 111–176. MIT Press.

Heim, I. & A. Kratzer (1998). *Semantics in generative grammar*. Blackwell.

Maier, E. (2021). Emojis as pictures. Ms.

Marantz, A. (1995). Cat as a phrasal idiom. Ms.

Potts, C. (2005). *The logic of conventional implicatures*. OUP.

Shan, C.-C. (2002). Monads for natural language semantics. *Proceedings of ESSLLI6*.

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

Song, C. (2021a). A typology of semilexicality and the locus of grammatical variation. Talk at ICFL9.

Song, C. (2021b). On the semantics of root syntax. *Proceedings of LENLS18*.

Song, C. (2022a). The logic of words. Talk at AAL2022.

Song, C. (2022b). Sentence-final particle vs. sentence-final emoji. Talk at Cambridge SyntaxLab.