AN EVOLUTIONARY APPROACH TO THE TYPOLOGY OF FUNCTIONAL EXPRESSIONS

THE LINGUISTIC CIRCLE OF COPENHAGEN
MARCH 10, 2021

Jürgen Bohnemeyer
OUTLINE

- A typological puzzle
- A theory of functional expressions
- Some data
- An evolutionary model of grammaticalization
- Supporting evidence: the ecology of definiteness marking
- Summary
A TYPOLOGICAL PUZZLE

- taking a page from Sapir (1921: 86-126)
- though no fowl shall be harmed in the present version
- all of the following utterances are responses to Scene 20 of Wilkins (2016)

```
20. Spkr and Addr are inside a house looking out of (open) door. They are near the doorway. The referent is just outside of door (near it). The referent is easily reached by both Addr and speaker (and equidistant from both). “I like ______ book/radio.”
   “Who’s book/radio is ___?”
   • Does it make a difference if the Spkr points or not? Must Spkr point?
   • Does it make a difference if object has been mentioned before?
   • Does it make a difference if Addr already has attention on object vs. attention being drawn?
   • Does term change with change in closeness of Spkr/Addr to door? Closeness of object to door?
```

Figure 1.1. Scene 20 of Wilkins (2016)
(1.1) Hijazi Arabic (Saudi Arabia)

Li-man ḏī:ḥ el-kita:b?
to-who this(SG) DEF-book
‘Whose is this book?’ (Ali M. Alshehri, p.c.)

(1.2) German

We-m gehör-t dies-es Buch?
who-DAT.SG belong-3SG.PRS this-SG.N.NOM book
‘To whom does this book belong?’

(1.3) Japanese (colloquial)

Kono hon dare=no?
this book who=GEN
‘Whose book is this?’ (Mitsuaki Shimojo, p.c.)

(1.4) Japanese (formal)

Kono hon (=wa) dare=no desu ka?
this book=TOP who=GEN COP Q
‘Whose is this book?’ (Mitsuaki Shimojo, p.c.)
(1.5) Saliba-Logea (Oceanic, Papuan Tip)

\textbf{Kaiteya yo-na tobwa ina?}

who CL1-3SG.POSS bag this

‘Whose bag is this?’ (Margetts 2016: 261)

(1.6) Yucatec Maya (Mexico)

\textbf{Máax ti’a’l le=liibro he’l=o’?}

who property(B3SG) DEF=book PRSV=D2

‘Whose property is that book over there?’

\textbf{Table 1.1. Summary of functional meanings expressed in (1.1)-(1.6)}

<table>
<thead>
<tr>
<th>Expressed in all examples</th>
<th>Expressed only in particular languages</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Interrogative pro-form</td>
<td>• Definiteness (Arabic, Yucatec)</td>
</tr>
<tr>
<td>• Demonstrative (or complex expression of spatial deixis)</td>
<td>• Case (German, Japanese)</td>
</tr>
<tr>
<td>• (Complex expression of) possession</td>
<td>• Gender/noun class (German, Saliba)</td>
</tr>
<tr>
<td>• Interrogative speech act (mostly coded via sentence type construction)</td>
<td>• Number (Arabic, German, Saliba, Yucatec)</td>
</tr>
<tr>
<td></td>
<td>• Person (German, Saliba, Yucatec)</td>
</tr>
<tr>
<td></td>
<td>• Topic (Japanese)</td>
</tr>
</tbody>
</table>
what accounts for this differential distribution?

my answer in a nutshell

the functional meanings that are expressed in all languages are part of the speaker’s intended message

the typologically variable functional expressions serve to facilitate comprehension

Table 1.2. Communicative functions of constituents of (1.1)-(1.6)

<table>
<thead>
<tr>
<th>Speaker’s intended message</th>
<th>Redundant facilitative categorizations</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Draw ADDR’s attention to $x$</td>
<td>• Topic time includes utterance time</td>
</tr>
<tr>
<td>• Lexically categorize $x$</td>
<td>• Uniqueness of $x$ in speech situation</td>
</tr>
<tr>
<td>• Relate $x$ to possessor $y$</td>
<td>• $x$ is inanimate; $</td>
</tr>
<tr>
<td>• Ask about identity of $y$</td>
<td>• Possessed: $x$; possessor: $y$</td>
</tr>
<tr>
<td></td>
<td>• Topic: $x$; focus: $y$</td>
</tr>
</tbody>
</table>
but what mechanism causes non-redundant functional expressions to be distributed near-universally

- and redundant ones to be distributed much more variably?

- to account for this, we need an upgrade to grammaticalization theory (GT)

- a mechanism for functional selection

- that boosts the grammaticalization of expressions adapted for communicative fitness

Figure 1.2. Grammaticalization and functional selection
the bigger picture: typology and evolutionary linguistics


Figure 1.3. The interaction between typology and evolutionary linguistics
the organization of this talk
(and the organization of the book I’m working on)

Figure 1.4. Organization of the talk/book as a flowchart
commitments

- **evolutionary functionalism** - only evolutionary models can explain how semantic/pragmatic functions shape language as long as teleological explanations are rejected

- **constructionism** - constructions are simple or templatic signs with conventionalized iconic and symbolic meanings with morphophonological, morphosyntactic, semantic, and pragmatic properties that continuously evolve
cf. Langacker (1987); Goldberg (1995); Croft (2001); Boas & Sag (2012); *inter alia*
commitments (cont.)

- **probabilistic pragmatics** - comprehension is based on (often non-monotonic) inferences, not on decoding

  - cf. Grice (1975, 1989); Sperber & Wilson (1987); Clark (1996); Levinson (2000); Goodman & Frank (2016); *inter alia*

- **categorical particularism** - constructions and types of functional expressions are strictly language-specific = *emic*

  - the *etic* ‘comparative concepts’ we use to compare them have no explanatory value

  - cf. Dryer (1997); Haspelmath (2007, 2010); *inter alia*
OUTLINE

- A typological puzzle
- A theory of functional expressions
- Some data
- An evolutionary model of grammaticalization
- Supporting evidence: the ecology of definiteness marking
- Summary
A THEORY OF FUNCTIONAL EXPRESSIONS

- the role of functional expressions in language

Figure 2.1. The place of functional expressions in the grammar of language $L$
so what are functional expressions? In first approximation:

- functional expressions are part of the grammar of the language as *individual expressions* (but types, not tokens)
  - rather than as members of larger categories
  - that is, there are construction templates/rules that reference the individual functional expressions
    - e.g., in English
      - the preposition *of* in possessive constructions
      - the verb *be* in nonverbal predication and progressive aspect constructions
so what are functional expressions? (cont.)

this is not a new insight

“Roughly, then, the total stock of elementary forms of a language can be split into two unequal portions: tea, write, and all other grammatically ‘unimportant’ forms go into one portion (by far the larger), while he, she, and all other grammatically ‘important’ forms go into the other. The deletion of anyone or two forms from the first portion would leave the grammatical system of the language essentially unchanged; the deletion of even a single item of the second kind would have drastic consequences. Equally drastic consequences could not be achieved by tinkering with the first portion unless we deleted all the members of some large form-class” (Hockett 1958: 261-262).
on this view, most, but not all, functional expressions are

- closed-class items
- grammaticalized

- e.g., *gehören* in (2.1) is an ordinary verb
  and *ti‘a‘l* in (2.2) an ordinary noun

(2.1) German

```
We-m gehör-t dies-es Buch?
```

who-DAT.SG belong-3SG.PRS this-SG.N.NOM book

‘To whom does this book belong?’

(2.2) Yucatec Maya (Mexico)

```
Máax ti‘a‘l le=liibro he‘l=o’?
```

who property(B3SG) DEF=book PRSV=D2

‘Whose property is that book over there?’
variables that form the basis of the classification of functional expressions

- communicative function: discourse-prominent vs. inherently backgrounded
- combinatorial and semiotic properties
communicative function/discourse prominence

inspired by Boye & Harder (2012)

classifies functional expressions into those that may express \textit{at-issue content} and those that may not

the latter are said to be \textit{inherently backgrounded}
communicative function/discourse prominence (cont.)

- **at-issue content**: provides a (partial) answer to the context’s *question under discussion* (QuD)
  - by reducing the number of live alternatives that are consistent with the discourse
communicative function/discourse prominence (cont.)

- the QuD of an utterance’s context determines the utterance’s information perspective
  - provided the utterance is felicitous and the discourse coherent

(2.3) a. [Q: Who ate the cake? – A:] FLOYD (did/ate the cake).
   b. [Q: What did Floyd eat? – A:] (He ate) the CAKE.
   c. [Q: What did Floyd do to the cake? – A:] EAT it / He ATE it.
   d. [Q: What happened next? – A:] FLOYD ATE THE CAKE.
communicative function/discourse prominence (cont.)

- inherent backgrounding of functional expressions means they cannot express at-issue content
  - and thus cannot be focalized
  - e.g., the past tense marker in (2.4) cannot be focalized
    - stress on the auxiliary expresses ‘verum focus’
  - but the negation can

(2.4) Q: DID Floyd eat the cake? – A: NO.
communicative function/discourse prominence (cont.)

- Boye & Harder (2012) treat discourse-primary expressions as not grammaticalized and not part of the grammar.

- In contrast, the present approach allows for discourse-prominent functional expressions by treating discourse prominence as one of two (give or take) properties.

- That govern the classification of functional expressions.
combinatorial and semiotic properties

Cann (2000): functional categories can be defined in terms of language-specific distributional classes

vis-à-vis the major lexical categories V, N, A

Figure 2.2. Lattice representing a taxonomy of nominal functional categories of English defined in terms of distributional classes (Cann 2000: 18)
combinatorial and semiotic properties (cont.)

one way to visualize the semantic effect of combinatorial properties: semantic type

Table 2.1. *Standard-issue extensional Montegovian type system for English sans events/situations*
combinatorial and semiotic properties (cont.)

- the point: functional expressions differ from lexical expressions in that they are syncategoremata and more abstract and relational in their meanings

<table>
<thead>
<tr>
<th>Lexical and phrasal categories</th>
<th>Functional expressions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Proper nouns, pronouns</td>
<td>$e$</td>
</tr>
<tr>
<td>Non-relational common nouns, standard-form predicative adjectives, intransitive verbs, VPs</td>
<td>$&lt;e,t&gt;$</td>
</tr>
<tr>
<td>NPs headed by common nouns</td>
<td>$&lt;&lt;e,t&gt;,t&gt;$</td>
</tr>
<tr>
<td>Relational common nouns, comparative-form predicative adjectives, transitive verbs</td>
<td>$&lt;e,&lt;&lt;e,t&gt;&gt;$</td>
</tr>
<tr>
<td>Ditransitive verbs</td>
<td>$&lt;e,&lt;&lt;e,&lt;e,t&gt;&gt;&gt;$</td>
</tr>
<tr>
<td>Attributive adjectives, relative clauses</td>
<td>$&lt;&lt;e,t&gt;,&lt;e,t&gt;&gt;$</td>
</tr>
<tr>
<td>Clauses, sentences</td>
<td>$t$</td>
</tr>
</tbody>
</table>
combinatorial and semiotic properties (cont.)

consider also Sapir’s (1921: 68-128) classification of linguistic meanings

Table 2.2. Sapir’s (1921: 92-93) classification of concepts expressed in The farmer killed the duckling
combinatorial and semiotic properties (cont.)

- beyond semantic type, combinatorial properties in a broad sense also extend to
- indexicality
- the component of the meaning of the utterance that the functional expression operates on

Table 2.3. Hengeveld’s (1989: 131-132) classification of operators in Functional Grammar

<table>
<thead>
<tr>
<th>Semantic domain</th>
<th>Grammatical category</th>
</tr>
</thead>
<tbody>
<tr>
<td>Internal temporal constituency</td>
<td>Imperfective/Perfective, Phasal Aspect Predicate negation</td>
</tr>
<tr>
<td>Presence or absence of property or relation expressed by predicate</td>
<td></td>
</tr>
<tr>
<td>Time of occurrence</td>
<td>Tense</td>
</tr>
<tr>
<td>Frequency of occurrence</td>
<td>Quantificational Aspect</td>
</tr>
<tr>
<td>Actuality of occurrence</td>
<td>Objective mood/Polarity</td>
</tr>
<tr>
<td>Source of proposition</td>
<td>Evidential mood</td>
</tr>
<tr>
<td>Commitment to proposition</td>
<td>Subjective mood</td>
</tr>
<tr>
<td>Weakening strategy</td>
<td>Mitigating mode</td>
</tr>
<tr>
<td>Strengthening strategy</td>
<td>Reinforcing mode</td>
</tr>
</tbody>
</table>

Table 2.4. Operators in the layered structure of the clause (Van Valin 2005: 9)

Nuclear operators:
- Aspect
- Negation
- Directionals (only those modifying orientation of action or event without reference to participants)

Core operators:
- Directionals (only those expressing the orientation or motion of one participant with reference to another participant or to the speaker)
- Event quantification
- Modality (root modals, e.g. ability, permission, obligation)
- Internal (narrow scope) negation

Clausal operators:
- Status (epistemic modals, external negation)
- Tense
- Evidentials
- Illocutionary force
how communicative and combinatorial properties come together

some discourse-prominent expressions are lexical, others are part of the grammar
due to their semiotic and combinatorial properties
what unites them is that they are needed to express the speaker’s intended message

in contrast, inherently backgrounded functional expressions are redundant wrt. the intended message
but instead serve to boost the odds that the hearer will infer the intended message
how communicative and combinatorial properties come together (cont.)

- this distinction between communicatively primary and secondary functional expressions is not a new idea
  
  - and neither is the observation that the latter are typologically more variable than the former

“We are thus once more reminded of the distinction between essential or unavoidable relational concepts and the dispensable type. The former are universally expressed, the latter are but sparsely developed in some languages, elaborated with a bewildering exuberance in others.” (Sapir 1921: 99).
the classification: let’s have it!

**Figure 2.3.** The proposed classification of functional expressions
the classification: hybrids

there quite a few pervasive expressions in natural languages that instantiate multiple types at ones

e.g., morphologically unbound personal pronouns are placeholders (discourse-prominent)

but the co-expressed categories of person, number, gender are restrictors on them (inherently backgrounded)

similarly for the distance distinctions in demonstratives
the classification: limits

- the proposed theory contains no principles from which to derive an exhaustive classification of functional expressions
  - beyond the level of the seven super-types
- I’m skeptical that an exhaustive classification is possible
the classification: limits (cont.)

- it’s part of the “logic” of evolution
  
  - that the possibility space for the emergence of new species depends on the set of existing species
    
    - e.g., birds could not have evolved before dinosaurs
  
- thus the possibility space continues to evolve itself

- the languages that existed beyond the time horizon of the comparative method
  
  - probably actually were substantially different from today’s
predictions

Table 2.5. Predictions generated by the proposed theory of functional expressions

<table>
<thead>
<tr>
<th>Type of functional expression</th>
<th>Discourse-prominent</th>
<th>Inherently backgrounded</th>
</tr>
</thead>
<tbody>
<tr>
<td>Degree of grammaticalization</td>
<td>Weaker</td>
<td>Stronger</td>
</tr>
<tr>
<td>Typological distribution</td>
<td>Near-universal: expressed in any all-purpose language, either by a functional expression or compositionally</td>
<td>Variable to a degree that depends on just how redundant the type of expression is</td>
</tr>
<tr>
<td><em>Ex nihilo</em> type innovation (type emergence in a genus without contact model)</td>
<td>Limited to transitions b/w compositional and non-compositional expressions</td>
<td>Unlimited</td>
</tr>
</tbody>
</table>
OUTLINE

- A typological puzzle
- A theory of functional expressions
- Some data
- An evolutionary model of grammaticalization
- Supporting evidence: the ecology of definiteness marking
- Summary
SOME DATA

- discourse-prominent expressions

- demonstratives have been argued to be present in all languages (Diessel 1999; Dixon 2003)

- exceptions arise in languages that use compositional expressions for exophoric reference

- such as French and Yucatec

(3.1) Yucatec Maya (Mexico)
Máax ti’a’l le=liibro (he’l)=o’?
who property(B3SG) DEF=book PRSV=D2
‘Whose property is that book over there?’

Table 3.1. French demonstrative paradigms
(Diessel 1999: 37)
similarly, independent pronouns are present universally

though some languages have compositional pronoun stems

(3.1) Mundari (Mundar, India; Daniels 2013)

a-ṉ ‘I’  a-liŋ ‘we.DU’  a-le ‘we.PL’;
a-m ‘you.SG’ a-ben ‘you.DU’  a-pe ‘you.PL’

Everett (2005) argues that Pirahã had no independent pronouns before borrowing some from Tupian languages

Evans & Levinson (2009: 431) claim that

“Sign languages like ASL (American Sign Language) also lack pronouns, using pointing instead.”

Cormier et al. (2013) dispute this
data from the Atlas of Pidgin and Creole Language Structures (Michaelis et al. 2013)

all sample languages have expressions equivalent to

- demonstratives
  - complex circumnominal forms often but not always involve augmentation with adverbs (‘the/that N there’)
- independent pronouns
- interrogative pro-forms
- negations
- frequency adverbs
- cardinal numerals
- adpositions (defined purely syntactically)
- verbal and NP conjunctions (defined purely syntactically)
person and number distinctions are restrictors on pronouns

so it is not surprising that there are counterexamples to Greenberg’s Universal 42

“All languages have pronominal categories involving at least three persons and two numbers” (Greenberg 1966: 96).

e.g., Everett (2005) reports that Pirahã does not express number either in nouns or in pronouns

there are eight languages in the APiCS sample w/ 1/2 or 2/3 syncretism (Haspelmath 2013)

Cysouw (2009: 39-65) discusses additional examples

similarly, 14 of 75 sample languages lack distance distinctions in demonstratives
restrictors: APiCS and WALS

**Table 3.1.** Some restrictor types in the APiCS and WALS databases

<table>
<thead>
<tr>
<th>Type of functional expression</th>
<th>APiCS</th>
<th>WALS</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Sample languages</td>
<td>Number of sample languages the type is attested in</td>
</tr>
<tr>
<td>Case</td>
<td>76</td>
<td>15 (19.7%)</td>
</tr>
<tr>
<td>Subject agreement</td>
<td>75</td>
<td>45 (60%)</td>
</tr>
<tr>
<td>Nominal gender</td>
<td>No data</td>
<td>170</td>
</tr>
<tr>
<td>Nominal number</td>
<td>76</td>
<td>71 (93.4%)</td>
</tr>
<tr>
<td>Past tense</td>
<td>75</td>
<td>59 (78.7%)</td>
</tr>
<tr>
<td>Definite article</td>
<td>76</td>
<td>57 (75%)</td>
</tr>
</tbody>
</table>
the distributions of case marking and subject agreement/cross-reference are very roughly complementary

this is not surprising, as head marking and dependent marking are alternative strategies

but, this makes structural case and agreement an obvious example of overlapping/competing redundant categories

Figure 3.1. Distribution of clause-level argument marking strategies (Nichols & Bickel 2013)
“ex-nihilo” innovations

- innovation of functional expressions not inherited from the genealogical ancestor in the absence of a contact model
  - in practice, absence of the type of functional expression in question in the other members of the genus
    - usually has to serve as a stand-in for evidence of absence of genealogical transmission
- prediction: innovations of discourse-prominent functional expressions are limited to transitions
  - between compositional and non-compositional expressions
  - in contrast, ex-nihilo innovation of inherently backgrounded functional expressions ought to be more common
“ex-nihilo” innovations (cont.)

- Evidence of ex-nihilo innovations of functional expressions is key proof of concept for any evolutionary theory.
- Ex-nihilo innovations directly attest to the evolvability of the particular type of expression.
a few attested examples of *ex-nihilo* innovations of functional expressions

- Wälchli (2018) discusses the emergence of gender in Nalca (Mek, Tanah Papua)

- Egyptian likely innovated articles and structural case (Levin 1992; Eitan Grossman, p. c.)

- Gullah (creole, Carolinas and Georgia) has numeral classifiers (Mufwene 1986)
  - but neither the lexifier nor any of the likely substrate languages does

- Matthew Dryer (p. c.) reports the innovation of an imperative mood out of an irrealis mood
  - and a diminutive out of a neuter gender in Walman (Torricelli, PNG)
OUTLINE

- A typological puzzle
- A theory of functional expressions
- Some data
- An evolutionary model of grammaticalization
- Supporting evidence: the ecology of definiteness marking
- Summary
AN EVOLUTIONARY MODEL OF GRAMMATICALIZATION

traditional grammaticalization theory (GT)

grammar is the result of grammaticalization

grammaticalization involves unidirectional change along loosely correlated scales

of metaphoric extension, semantic bleaching, and morphophonological reduction

Lehmann 1982; Heine & Reh 1984; Hopper 1981; inter alia

Table 4.1. Grammaticalization processes (Croft 2000: 157)

<table>
<thead>
<tr>
<th>Phonological</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Paradigmatic:</td>
<td>attrition: reduction/erosion &gt; phonological loss</td>
</tr>
<tr>
<td>Syntagmatic:</td>
<td>coalescence: free morpheme &gt; cliticization, compounding &gt; affixation &gt; loss adaptation (including assimilation)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Morphosyntactic</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Paradigmatic:</td>
<td>obligatorification &gt; fossilization &gt; morphological loss paradigmaticization: open class &gt; closed class &gt; invariant element</td>
</tr>
<tr>
<td>Syntagmatic:</td>
<td>rigidification [word order]</td>
</tr>
<tr>
<td></td>
<td>loss of independent syntactic status &gt; morphological fusion &gt; loss</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Functional</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Paradigmatic:</td>
<td>extension of semantic range &gt; loss of function</td>
</tr>
<tr>
<td>Syntagmatic:</td>
<td>idiomaticization: compositional &amp; analyzable &gt; noncompositional &amp; analyzable &gt; unanalyzable</td>
</tr>
</tbody>
</table>
GT takes a “ballistic” view of grammaticalization

- the output of grammaticalization processes is seen as constrained only by the input
  - and by *mutational constraints* (Hasnelmath 1999)
    - i.e., by the available pathways of grammaticalization and the unidirectionality of grammaticalization
- mutational constraints alone cannot account for the differential typological distribution of functional expressions
- the sources of the grammaticalization of inherently backgrounded expressions are lexical expressions and discourse-prominent functional expressions
- since both of these are universally available, why would backgrounded functional expressions not also be?

Figure 4.1. Some grammaticalization pathways in the verbal domain (Lehmann 2015: 39)
to account for the observable typological distribution of functional expressions

- GT must be retrofitted with **functional-adaptive constraints** (Haspelmath 2019)
  - i.e., constraints that boost the grammaticalization of expressions that are optimized for communicative fitness

- Speakers have a choice from among competing structural alternatives for communicating the same message.
- Selection among these is biased in terms of 'user optimality', i.e., communicative fitness.
- This boosts the usage frequency of the fitter options, causing regularization and obligatorization.
- And potentially eventually the loss of the competitors.
a variety of different evolutionary models of language change have been proposed recently

Table 4.2. A typology of evolutionary models of language change

<table>
<thead>
<tr>
<th>Specifity of the model</th>
<th>Assumed primary “agents” of language change</th>
<th>(Primarily L1) Learners</th>
<th>(Primarily adult) speakers</th>
<th>Both learners and speakers</th>
</tr>
</thead>
</table>
sketching an explicit causal model

Assume a contrast between two expressions $C_1$ and $C_2 = C_1 + x$, both of which could be used to express the speaker’s communicative intent. The addition of $x$ to $C_1$ is licensed by an existing construction of the language. Its use in $C_2$ is redundant wrt. the speaker’s communicative intent, but increases the probability $p$ of the hearer inferring the intended meaning $I$ ($p(I|C_2) > p(I|C_1)$). E.g., $x$ could be a demonstrative added to indicate definiteness, a perfective aspect marker to indicate past time reference, or an allative case marker to indicate a syntactic object relation.

**Figure 4.2.** An evolutionary model of the grammaticalization of redundant functional expressions
OUTLINE

- A typological puzzle
- A theory of functional expressions
- Some data
- Non-evolutionary explanations
- An evolutionary model of grammaticalization
- Supporting evidence: the ecology of definiteness marking
- Summary
SUPPORTING EVIDENCE: THE ECOLOGY OF DEFINITENESS MARKING

- Evers (2020): rationale
  - if the grammaticalization of inherently backgrounded functional expressions is a functional adaption
    - it should occur where there are functional “niches” for it
      - and not elsewhere
  - so it should be possible to predict which languages grammaticalize e.g. definite articles and which don’t
    - on the basis of the presence/absence of alternative morphosyntactic definiteness cues
grammar sampling study: generating the sample

- a sample of 100 languages were randomly selected by an algorithm introduced in Dryer (2018)

- based on two criteria
  - availability of a recent extensive description
  - at least 30 languages spoken geographically in between each pair of adjacent languages

Figure 5.1. Map of the language sample of Evers (2020: 125)
grammar sampling study: procedure

- manually coding the sample languages for 8 variables selected from an original 16 after a pilot study

- on a sample of 32 languages at a distance of 50 languages in between

- run machine learning models to identify the strongest predictors of absence of definite articles

Figure 5.2. Pilot (left) and final set of independent variables (Evers 2020: 88, 126)
grammar sampling study: findings

**Figure 5.3.** Random forest models of the main sample predicting absence of definite articles (Evers 2020: 135)

**Figure 5.4.** Conditional inference tree of the main sample predicting absence of definite articles (Evers 2020: 136)
Evers followed this analysis up with two corpus studies on Kalaallisut (Eskaleut, Greenland) and Colloquial Jakarta Indonesian (CJI), both of which lack definite articles. She manually coded discourses in both languages and ran classifiers predicting definiteness, and she found that models accurately predicted definiteness in 78% of arguments in CJI and 90% of arguments in Kalaallisut.
OUTLINE

- A typological puzzle
- A theory of functional expressions
- Some data
- Non-evolutionary explanations
- An evolutionary model of grammaticalization
- Supporting evidence: the ecology of definiteness marking
- Summary
there are two types of functional expressions in the languages of the world

- discourse-prominent expressions are capable of expressing at-issue content
  - i.e., part of the speaker’s intended message
- they share this property with lexical expressions
  - but differ from them in terms of their combinatorial properties and abstract, syncategorematic semantics
in contrast, inherently backgrounded functional expressions are communicatively redundant to varying degrees

- their primary purpose is to boost the odds that the hearer will infer the intended meaning
- their function is thus primarily metalinguistic and they tend to be more strongly grammaticalized
discourse-prominent functional expressions are distributed near-universally across the languages of the world

- every “all-purpose” language expresses the relevant meanings either compositionally or non-compositionally

- in contrast, inherently backgrounded functional expressions display considerable typological variation

- and it appears that the extent of this variation correlates with the extent of their backgrounding/redundancy

- evidence from ex-nihilo innovations further supports this conclusion

- and represents direct evidence of the evolvability of functional expressions
theories of grammaticalization that take into account only sources and mutational constraints

cannot explain the typological distribution of functional expressions

to do this, grammaticalization theory requires an upgrade with an evolutionary module of functional selection
evidence in support of the idea that inherently backgrounded functional expressions evolve where they fill functional niches.

comes from a typological study using machine learning models to predict the absence of definite articles on the basis of competing alternative definiteness cues.
OUTLINE

- Emotions
- Heider on interpersonal emotions
- Force dynamics of emotion
- Semiotics of emotion in language
- Emotion across cultures
ACKNOWLEDGMENTS

- My thinking about the material presented here has greatly benefitted from input and feedback by
  - Kasper Boye, Lea Brown, Östen Dahl, Matthew Dryer, Stephanie Evers, David Gil, Jeff Good, Eitan Grossman, Peter Harder, Martin Haspelmath, Christian Lehmann, Johanna Nichols, Adam Tallman, Robert Van Valin Jr., Samira Verhees, Berhard Wälchli; *inter alia*
  - the participants of the RRG 2019 conference
  - the members of the UB Semantic Typology Lab
- None but me should be presumed to agree with, or be responsible for, any of the material
THANKS!