

AN EVOLUTIONARY APPROACH TO THE TYPOLOGY OF FUNCTIONAL EXPRESSIONS

LANGUAGE CULTURE COGNITION LAB AUSTRIAN ACADEMY OF SCIENCES NOVEMBER 30, 2023

Jürgen Bohnemeyer **University at Buffalo** *The State University of New York*

OUTLINE

- A typological puzzle
- A theory of functional expressions
- A smattering of data
- Studying the distribution of functional expressions
- An evolutionary model of grammaticalization
- Supporting evidence: the ecology of definiteness marking
- Summary
- Appendices

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 [1999])

20.



Figure 1.1. Scene 20 of Wilkins (2016)

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?

(1.1) Hijazi Arabic (Saudi Arabia)

Li-man ði:h el-kita:b? to-who this(SG) DEF-book 'Whose is this book?' (Ali M. Alshehri, p.c.)

(1.2) German

We-mgehör-tdies-esBuch?who-DAT.SGbelong-3SG.PRSthis-SG.N.NOMbook'To whom does this book belong?'

(1.3) Japanese (colloquial)

Kono hondare=no?thisbookwho=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)
 Kaiteya yo-na tobwa ina?
 who CL1-3SG.POSS bag this
 'Whose bag is this?' (Margetts 2016: 261)

(1.6) Yucatec Maya (Mexico)

Máaxti'a'lle=lìibrohe'l=o'?whoproperty(B3SG)DEF=bookPRSV=D2'Whose property is that book over there?'

Table 1.1. Summary of functional meanings expressed in (1.1)-(1.6)

Expressed in all examples Expressed only in particular languages	
• Interrogative pro-form	• Definiteness (Arabic, Yucatec)
• Demonstrative (or complex	• Case (German, Japanese)
expression of spatial deixis)	• Gender/noun class (German, Saliba)
• (Complex expression of)	• Number (Arabic, German, Saliba,
possession	Yucatec)
• Interrogative speech act	• Person (German, Saliba, Yucatec)
(mostly coded via sentence	• Topic (Japanese)
type construction)	

- 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)

eaker's intended message Redundant facilitative categorizations	
• Draw ADDR's attention to x	• Topic time includes utterance time
• Lexically categorize x	• Uniqueness of x in speech situation
• Relate x to possessor y	• \underline{x} is inanimate; $ x = 1$
• Ask about identity of y	• Possessed: x ; possessor: y
	• Topic: x ; focus: y

- 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 to upgrade
 grammaticalization theory (GT)
 - with a mechanism for functional selection
 - that boosts the grammaticalization of expressions adapted for communicative fitness



Figure 1.2. Grammaticalization and functional selection

OUTLINE

- A typological puzzle
- A theory of functional expressions
- A smattering of data
- Studying the distribution of functional expressions
- An evolutionary model of grammaticalization
- Supporting evidence: the ecology of definiteness marking
- Summary
- Appendices

A THEORY OF FUNCTIONAL EXPRESSIONS

- so what are functional expressions?
 - morphemes that are part of the grammar of the language as individual expressions (but as types, not tokens)
 - rather than as members of lexical/syntactic 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 then spells out the usual suspects
 - function words
 - inflections
 - highly productive and transparent derivations

- 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-mgehör-tdies-esBuch?who-DAT.SGbelong-3SG.PRS this-SG.N.NOMbook'To whom does this book belong?'

(2.2) Yucatec Maya (Mexico)

Máax ti'a'lle=lìibrohe'l=o'?whoproperty(B3SG)DEF=bookPRSV=D2'Whose property is that book over there?'

- variables that form the basis of the classification of functional expressions
 - combinatorial and semiotic properties
 - ★ semantic type (Montague 1970, 1973; Klein & Sag 1985; Cann 2000)^{the, a, some, we,, few}



- communicative function/discourse prominence
 - inspired by Boye & Harder (2012)
 - classifies functional expressions into those that may express at-issue content and those that may not
 - the latter are said to be 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
 - Carlson (1982), Klein & von Stutterheim (1987, 2002), van Kuppevelt (1995, 1996), Roberts (1996, 2012), Büring (1997, 2003)

- 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 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
 - including demonstratives, pronouns, modals, etc.
 - 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

- 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



positional moods

Figure 2.2. Classification of functional expressions in terms of communicative import and semantic type



- 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).

OUTLINE

- A typological puzzle
- A theory of functional expressions
- A smattering of data
- Studying the distribution of functional expressions
- An evolutionary model of grammaticalization
- Supporting evidence: the ecology of definiteness marking
- Summary
- Appendices

A SMATTERING OF 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=**lìibro (he'l)**=o'**? who property(B3SG) DEF=book PRSV=D2 'Whose property is that book over there?'

Table 3.1. French demonstrative paradign (Diessel 1999: 37)

	DEM PROS		DEM DETS	
	PROXIMAL	DISTAL	PROXIMAL	DISTAL
G.M	celui-ci	celui-là	ce livre-ci	ce livre-là
G.F	celle-ci	celle-là	cette maison-ci	cette maison-là
L.M	ceux-ci	ceux-là	ces livres-ci	ces livres-là
L.F	celles-ci	celles-là	ces maisons-ci	ces maisons-là

- discourse-prominent expressions (cont.)
 - similarly, independent pronouns are present universally
 - though some languages
 have compositional pronoun stems
- (3.2) Mundari (Mundar, India; Daniels 2013)
 a-ñ 'l' a-li**ŋ** '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

A SMATTERING OF DATA (CONT.)

- discourse-prominent expressions (cont.)
 - all APiCS 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)

A SMATTERING OF DATA (CONT.)

- discourse-prominent expressions (cont.)
 - 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

Table 3.2. Some restrictor types in the APiCS and WALS databases



Figure 3.1. Some restrictor types in the APiCS and WALS databases

- "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 SMATTERING OF DATA (CONT.)

- "ex-nihilo" innovations (cont.)
 - 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
- A smattering of data
- Studying the distribution of functional expressions
- An evolutionary model of grammaticalization
- Supporting evidence: the ecology of definiteness marking
- Summary
- Appendices

STUDYING THE DISTRIBUTION OF FUNCTIONAL EXPRESSIONS

- what is the probability of a given type of functional expression C being present in a particular language L?
- I argue that the answer depends on
 - the phylogeny and areality (geography) of L
 - the informativeness/redundancy of C
 - the pragmatic resource ecology of *L*
 - i.e., the extent to which L has grammaticalized pragmatic competitors of C

STUDYING THE DISTRIBUTION OF FUNCTIONAL EXPRESSIONS (CONT.)

the role of phylogeny





Figure 4.1. The role of phylogeny in shaping the typological distribution of functional expressions: five scenarios. Key: Solid orange discs represent a feature of interest (here: a particular type of functional expression)

the role of contact/areality



Figure 4.2. The role of areality in shaping the typological distribution of functional expressions: two scenarios. Key: Solid orange discs represent a feature of interest (here: a particular type of functional expression); dotted line encircles a linguistic area.

STUDYING THE DISTRIBUTION OF FUNCTIONAL EXPRESSIONS (CONT.)

- correcting for phylogenetic and areal biases: APiCS
 - pidgins and creoles are new languages that are in theory free of phylogenetic biases
 - problems
 - possible lexifier and substrate influences
 - 58/76 APiCS languages have European lexifiers!
 - possible contact influences, including areal ones
 - cf. Blasi et al. (2017)
 - some pidgins are only used as trade/auxiliary languages



Percentage attested
STUDYING THE DISTRIBUTION OF FUNCTIONAL EXPRESSIONS (CONT.)

- correcting for phylogenetic and areal biases: stratified sampling
 - e.g. based on genera (Dryer 1989, 1992);
 geographic distance (Dryer 2018; Evers 2020)
 - strength: conservative
 - much more likely to produce a false negative
 - than a false positive
 - weakness: we don't know how to optimally stratify



Percentage attested

w/o knowing the size of phylogenetic and areal effects

- correcting for phylogenetic and areal biases: an alternative capitalizing on Grambank (Skirgård et al. 2023)
 - Grambank comes much closer than WALS to being free of missing observations
 - making regression modeling statistically feasible



Figure 4.4. Sample languages × variables in Grambank vs. WALS (Skirgård et al. 2023 Supplementary materials (DOI: 10.1126/sciadv.adg6175)

- correcting for phylogenetic and areal biases: an alternative capitalizing on Grambank (Skirgård et al. 2023) (cont.)
 - for each type of functional expression C in Grambank (GB)
 - compute for each phylogeny P the proportion p_{CP}
 of members of P in which C is grammaticalized
 - treat isolates and creoles as phylogenies
 - average the result across phylogenies





STUDYING THE DISTRIBUTION OF FUNCTIONAL EXPRESSIONS (CONT.)

- correcting for phylogenetic and areal biases: an alternative capitalizing on Grambank (Skirgård et al. 2023) (cont.)
 - using regression modeling
 to estimate the effects of areal and phylogenetic biases

```
C_L \sim \beta_0 + \beta_1 P_L + \beta_2 A_L + \beta_3 \bar{\boldsymbol{p}}_C
```

- C_L The probability of an expression of type C being grammaticalized in L
- P_L The phylogeny of L

(4.1)

*A*_L - The geographic coordinates of *L*

 \bar{p}_{CP} - The mean frequency of grammaticalization of C across phylogenies

- report **four** numbers to capture the typological distribution of each type of functional expression *C*
 - the total number of phylogenies in which C is grammaticalized
 - \bar{p}_C the mean by-phylogeny frequency of C
 - β_1 and β_2 the estimated extent of areal and phylogenetic biases (for those C for which P and/or A are significant)

- correcting for phylogenetic and areal biases: an alternative capitalizing on Grambank (Skirgård et al. 2023) (cont.)
 - possibly in addition generate world trees showing ancestral state estimates for each functional expression?
 - using e.g. INLA
 (Martins et al. 2013)

Figure 4.5. INLA-generated tree plot of GB feature 74, Are there prepositions? "Tip point colors represent observed values: black = yes (there are prepositions), uncolored = no (there are not prepositions), gray = missing data. Branch colors represent probability estimates: yellow = higher probability that there are prepositions, purple = lower probability that there are prepositions." (Skirgård et al. 2023: Supplementary Figure 5.)



OUTLINE

- A typological puzzle
- A theory of functional expressions
- A smattering of data
- Studying the distribution of functional expressions
- An evolutionary model of grammaticalization
- Supporting evidence: the ecology of definiteness marking
- Summary
- Appendices

- grammaticalization theory (GT)
 - grammar is the result of grammaticalization
 - grammaticalization involves unidirectional change along loosely correlated scales
 - of metaphoric/metonymic extension, semantic bleaching
 - and morphophonological reduction
 - Lehmann 1982; Heine & Reh 1984; Hopper 1981; Hopper & Traugott 1993; inter alia

Table 5.1. Grammaticalization processes(Lehmann 1985: 309)

parameter	weak grammaticalization	$-$ process \rightarrow	strong grammaticalization
integrity	bundle of semantic features; possibly polysyllabic	$-$ attrition \rightarrow	few semantic features; oligo- or monosegmental
paradigmaticity	item participates loosely in semantic field	$-$ paradigmaticization \rightarrow	small, tightly integrated paradigm
paradigmatic variability	free choice of items according to communicative intentions	$-$ obligatorification \rightarrow	choice systematically constrained, use largely obligatory
scope	item relates to constituent of arbitrary complexity	$-$ condensation \rightarrow	item modifies word or stem
bondedness	item is independently juxtaposed	$-$ coalescence \rightarrow	item is affix or even phonological feature of carrier
syntagmatic variability	item can be shifted around freely	$-$ fixation \rightarrow	item occupies fixed slot

- GT postulates that grammaticalization happens
 - and aims to provide a unified account of *how* it happens
 - i.e., an account that identifies the various processes that potentially or necessarily accompany grammaticalization
- however, GT does not inherently explain
 why grammaticalization happens

- why does grammaticalization happen? Some proposals
 - Lehmann (1985: 10-11): creativity (also, Haspelmath 1999b: 'extravagance')
 - the more strongly grammaticalized an expression
 - the less it satisfies the speaker's desire "to give the fullest possible expression" to her intended meaning
 - hence, speakers constantly creatively renew existing constructions with less grammaticalized ones



- but why do constructions tend to become more strongly grammaticalized over time, creating the need for creative renewal in the first place?
- and how do genuinely new constructions arise that don't have a model in the language in question?

- why does grammaticalization happen? Some proposals (cont.)
 - Heine et al. (1991: 27-32): communicative need
 - some concepts are inherently more grammatical than others
 - expressions of less grammatical concepts are extended to expressing more grammatical ones
 - through metaphor and metonymy
 - in processes that resembles exaptations in biology
 - these processes represent creative solutions to communicative needs
 - but what is the causal efficacy of communicative need?

- why does grammaticalization happen? Some proposals (cont.)
 - Boye & Harder (2012) (cf. also Keizer 2007)
 - grammaticalization changes expressions from being 'discourse-primary' to being 'discourse-secondary'
 - i.e., inherently backgrounded and therefore incapable of addressing the QuD (though B&H don't use 'QuD')
 - attrition and coalescence follow from thence
 - this explains some of the properties of grammaticalization
 - but it remains open what causes the creation of inherently backgrounded expressions

- why does grammaticalization happen? Some proposals (cont.)
 - Hopper & Traugott (2003 [1993]: 71-74); Geurts (2000) (cf. also Zipf 1949): efficiency and informativeness
 - competing motivations:
 - maximization of efficiency via minimal differentiation vs. maximization of informativeness
 - speakers strive to reduce the speech signal via routinization and signal simplification
 - simultaneously, they seek to boost their expressivity
 - thereby improving informativeness and simultaneously indexing social identities
 - but how do these competing forces effect changes that propagate through the speech community?

interlude: communicative efficiency and grammaticalization

"The Principle of Communicative Efficiency: Communicate in such a way as to maximize the benefit-to-cost ratio." (Levshina 2018:4)

- suppose the speaker's intended meaning in producing utterance U is situation s
- the speaker estimates the probability p of the hearer inferring s if U contains expression C: p(s|C)
 - Tishby et al. (1999); Frank & Goodman (2012); Franke & Jäger (2016); Kemp et al. (2018); *inter alia*
- the communicative efficiency of C in U is proportional to p and inversely proportional to the effort involved in uttering C

Interlude: communicative efficiency and grammaticalization



Figure 5.2. How grammaticalization (hypothetically) boosts communicative efficiency: Attempting a synthesis

- why does grammaticalization happen? Some proposals (cont.)
 - Croft (2000): evolution (cf. also Beckner et al. 2009)
 - the principal goal of evolutionary models
 - explaining adaptive change through ("natural") selection
 - i.e., without invoking teleological causation



Figure 5.3. Selection, adaption, and population dynamics

(https://flexbooks.ck12.org/cbook/ck-12-middle-school-earth-science-flexbook-2.0/section/17.4/primary/ lesson/adaptation-and-evolution-of-populations-ms-es/)

- why does grammaticalization happen? Some proposals (cont.)
 - Croft (2000): evolution (cf. also Beckner et al. 2009)
 - an evolutionary model of language change based on narrow biological analogies
 - selection of a new variant is its success in diffusion and transmission

 Table 5.2. Croft's (2000) extrapolation
 of Hull's (1988) model to Language change

	Biological evolution	Linguistic evolution
Replicator	Gene	Lingueme – a structural
		feature of an utterance (a
		word, construction,
		phonetic realization of a
		phoneme)
Interactor	Organism	Speaker
Selection	Biological procreation vs.	A lingueme (not) catching
	extinction	on (i.e., (not) being
		$\operatorname{transmitted}/\operatorname{diffused})$
Lineage	Lineages of biological	Etymologies; histories of
	descent	sounds/constructions

Croft argues that selection is driven solely by social factors such as overt/covert prestige and social relations

- why does grammaticalization happen? Some proposals (cont.)
 - Croft (2000): evolution (cont.)
 - communicative efficiency plays out solely in the minds of individual speakers
 - in the form of *altered replications* that may be biased toward enhanced communicative efficiency

- why does grammaticalization happen? Some proposals (cont.)
 - Croft (2000): evolution (cont.)
 - on Croft's account, the propagation of grammaticalization would be independent of its communicative fitness
 - but sociolinguistic forces are not sensitive to communicative fitness either
 - so changes of decreasing fitness should be just as likely to propagate as changes of increasing fitness
 - this seems to be at odds with the much-discussed (near-)unidirectionality of grammaticalization
 - cf. Lehmann (2015 [1982]: 18-21); Bybee et al. (1994: 12-14); Hopper & Traugott (2003 [1993]: 99-139); Haspelmath (1999b)

- why does grammaticalization happen? Some proposals (cont.)
 - Haspelmath (1999a: 190) (cf. also Hawkins 2014: 85-89): conditions on a functional-adaptive mechanism
 - based on Keller (1994) and Lüdtke (1980)

Conditions on a functional-adaptive mechanism in grammaticalization

- existence of structural variation:
 speakers have a choice
- the choice is influenced by communicative fitness
- feedback loop: biased choice influences frequency;
 frequency influences grammaticalization

sketching an explicit causal model

Conditions on a functional-adaptive mechanism in grammaticalization

- existence of structural variation: speakers have a choice
- the choice is influenced by communicative fitness
- feedback loop: biased choice influences frequency; frequency influences grammaticalization



Figure 5.4. An evolutionary model of grammaticalization

sketching an explicit causal model of the grammaticalization of redundant cue support

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.



the impact of system ecologies: an example

Iet

- L: the language in question
- $C_2 = [x C_1]$: expression C_1 augmented by redundant cue x
 - (e.g., C₁ is a nominal and x a demonstrative or pronoun used for optional definite marking)
- C₃: a potential existing pragmatically overlapping cue
 - (e.g., an alternative definiteness cue, such as (flexible) word order)
- if L has a C_3 , $p(I|C_2) p(I|C_1)$ is lower than if L lacks C_3
- hence, the presence of C_3 reduces the probability that speakers will use x for optional definiteness marking
 - and thus the odds of x grammaticalizing (e.g., into a definite article)

OUTLINE

- A typological puzzle
- A theory of functional expressions
- A smattering of data
- Studying the distribution of functional expressions
- An evolutionary model of grammaticalization
- Supporting evidence: the ecology of definiteness marking
- Summary
- Appendices

SUPPORTING EVIDENCE: THE ECOLOGY OF DEFINITENESS MARKING

- initial evidence that grammaticalization is indeed functionally adaptive comes from apparent tradeoffs
 - in what languages grammaticalize examples:
 - tradeoff between syntactic and morphological expression
 - Nichols (1992); Siewierska (1998); Koplenig et al. (2017)
 - John Shuey (ongoing research): near-complementary typological distribution
 - between syntactic causatives and highly productive morphological causatives
 - tradeoffs in the grammaticalization of tense, aspect, mood (Bhat 1999; Bohnemeyer 2000)

- Evers (2021): investigating tradeoffs in the grammaticalization of definiteness
- 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 6.1. *Map of the language sample of Evers (2021: 125)*



SUPPORTING EVIDENCE: THE ECOLOGY OF DEFINITENESS MARKING (CONT.)

- 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

Case	Differential Object Marking	
Gender	Noun Classifiers	
Perfective Aspect Marking	Switch Reference	
Prodrop	Verb Agreement	
Isolating	Ergativity	
Focus Marking	Topic Marking	
Flexible Subject Order	VO Word Order	
Affixation Order	Indefinite Article	

Alignment	Case Marking	
Morphological Complexity	Object Agreement	
Order of Object and Verb	Subject Order Flexibility	
Switch Reference	Topic Marking	

Figure 6.2. Pilot (left) and final set of independent variables (Evers 2021: 88, 126)

grammar sampling study: findings



Random Forest Results for Full Database Study



Definite Article Predictors

Figure 6.3. Random forest models of the main sample predicting **absence** of definite articles (Evers 2021: 135)

Figure 6.4. Conditional inference tree of the main sample predicting **absence** of definite articles (Evers 2021: 136)

SUPPORTING EVIDENCE: THE ECOLOGY OF DEFINITENESS MARKING (CONT.)

- 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
 - she found that models accurately predicted definiteness in
 - 78% of arguments in CJI
 - 90% of arguments in Kalaallisut

SUPPORTING EVIDENCE: THE ECOLOGY OF DEFINITENESS MARKING (CONT.)

- additional evidence: event order in discourse (Bohnemeyer 1998, 2000, 2002, 2009)
 - Yucatec Maya lacks both tense marking
 - and specific temporal connectives
 with meanings such as 'after' and 'before'
 - speakers are able to infer the order of events in discourse on the basis of aspect-mood marking
 - and conversational implicatures

- additional evidence: event order in discourse (cont.)
 - in a referential communication task, Yucatec speakers were as successful in communicating contrastive event orders

67

- as were the German-speaking control group
 - German having tense and specific temporal connectives, but only rudimentary aspect marking

OUTLINE

- A typological puzzle
- A theory of functional expressions
- A smattering of data
- Studying the distribution of functional expressions
- An evolutionary model of grammaticalization
- Supporting evidence: the ecology of definiteness marking

Summary

Appendices

SUMMARY

- across languages, functional expressions fall along a continuum of informativeness/redundancy
 - 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

- an evolutionary model of grammaticalization requires a functional-adaptive mechanism
 - key components of such a mechanism
 - speakers compare competing expressions of the same broad meaning in terms of communicative efficiency
 - their selections feed into the exemplar cloud
 - the rising frequency of communicatively fitter expressions causes their reduction and regularization
 - which in turn increases their fitness further
- 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 basic of competing alternative definiteness cues

ACKNOWLEDGMENTS

- My thinking about the material presented here has greatly benefitted from input and feedback by
 - Kasper Boye, Lea Brown, Bill Croft, Östen Dahl, Matthew Dryer, Stephanie Evers, David Gil, Jeff Good, Russell Gray, Eitan Grossman, Peter Harder, Martin Haspelmath, Nikolaus Himmelmann, T. Florian Jaeger, Christian Lehmann, Johanna Nichols, Peter Juul Nielsen, Hedvig Skirgård, Kenny Smith; Adam Tallman, Robert Van Valin Jr., Samira Verhees, Bernhard Wälchli; inter alia
 - the participants of the RRG 2019 conference and the 2023 FRIAS symposium Efficiency in Grammar
 - the members of the UB Semantic Typology Lab
 - the participants of my Spring 2021 seminar on Typology and Evolution
 - the people who attended my presentations of this material to the Linguistic Circle of Copenhagen and the University of Rochester audience design group
- None but me should be presumed to agree with, or be responsible for, any of the material



REFERENCES

- Aylett, M. & A. Turk. (2004). The Smooth Signal Redundancy Hypothesis: A functional explanation for relationships between redundancy, prosodic prominence, and duration in spontaneous speech. Language And Speech 47(1): 31-56.
- Beckner, C., Blythe, R., Bybee, J., Christiansen, M. H., Croft, W., Ellis, N. C., Holland, J., Ke, J., Larsen-Freeman, D., and Schoenemann, T. (2009). Language is a complex adaptive system: Position paper. Language Learning 59: 1–26.
- Bhat, D. N. S. (1999). The prominence of tense, aspect, and mood. Amsterdam: Benjamins.
- Blasi, D., S. M. Michaelis, & M. Haspelmath. (2017). Grammars are robustly transmitted even during the emergence of creole languages. Nature Human Behavior 1: 723-729.
- Bohnemeyer, J. (2000). Event order in language and cognition. In H. de Hoop and T. van der Wouden (eds.), Linguistics in the Netherlands 17. Amsterdam: John Benjamins. 1-16.
- Boye, K. & P. Harder. (2012). A usage-based theory of grammaticalization and grammatical status. Language 88(1): 1-44.
- Buz, E., M. K. Tanenhaus, & T. F. Jaeger. (2016). Dynamically adapted context-specific hyper-articulation: Feedback from interlocutors affects speakers' subsequent pronunciations. Journal of Memory and Language 89: 68-86.
- Bybee, J. (2007). Frequency of use and the organization of language. Oxford: Oxford University Press.
- Bybee, J., R. D. Perkins, & W. Pagliuca. (1994). The evolution of grammar: tense, aspect and modality in the languages of the World. Chicago: University of Chicago Press.
- Cann, R. (2000). Functional versus lexical: A cognitive dichotomy. In R. Borsley (ed.), The nature and function of syntactic categories. Syntax and Semantics 32: 37-78.
- Croft, W. (2000). Explaining language change: An evolutionary approach. Harlow: Pearson Education Limited.
- Dye, M. et al. (2018) Alternative solutions to a language design problem: the role of adjectives and gender marking in efficient communication. Topics Cognit. Sci. 10, 209-224.
- Evers, S. (2021). To 'the' or not to 'the': Cross-linguistic correlations between existing morphosyntax and the emergence of definite articles. Doctoral dissertation, University at Buffalo.
- Fedzechkina, M., T. F. Jaeger, & E. L. Newport. (2012). Language learners restructure their input to facilitate efficient communication. PNAS 109(44): 17897-17902.
- Fedzechkina, M. & T. F. Jaeger. (2020). Production efficiency can cause grammatical change: Learners deviate from the input to better balance efficiency against robust message transmission. Cognition 196: article 104115.

REFERENCES (CONT.)

- Fenk-Oczlon, G., & A. Fenk. (1999). Cognition, quantitative linguistics, and holistic typology. Linguistic Typology 3(2): 151-177.
- Ferrer-i-Cancho, R. (2017). The placement of the head that maximizes predictability: An information theoretic approach. Glottometrics 39: 38-71.
- Futrell, R., Mahowald, K., & Gibson, E. (2015). Large-scale evidence of dependency length minimization in 37 languages. Proceedings of the National Academy of Sciences 112: 10336–10341.
- Genzel, D. & E. Charniak. (2002). Entropy rate constancy in text. Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics 199-206.
- Geurts, B. (2000). Function or fashion? Reply to Martin Haspelmath. Linguistics 38(6): 1175-1180.
- Gibson, E. (1991). A computational theory of human linguistic processing: Memory limitations and processing breakdown. PhD thesis, Carnegie Mellon Univ.
- Haspelmath, M. (1999a). Optimality and diachronic adaptation. Zeitschrift für Sprachwissenschaft 18(2): 180-205.
- Haspelmath, M. (1999b). Why is grammaticalization irreversible? Linguistics 37(6): 1043-1068.
- Hawkins, J. A. (2014). Cross-linguistic variation and efficiency. Oxford: Oxford University Press.
- Heine, B., U. Claudi, & F. Hünnemeyer. (1991). Grammaticalization: A conceptual framework. Chicago: University
- of Chicago Press.
- Heine, B. & K.-A. Song. (2011). On the grammaticalization of personal pronouns. Journal of Linguistics 47(3): 587-630.
- Hopper, P. J. & E. C. Traugott. (20032 [1993]). Grammaticalization. Cambridge: Cambridge University Press.
- > Jaeger, T. F. (2010). Redundancy and reduction: Speakers manage syntactic information density. Cognitive Psychology 61(1): 23-62.
- > Jespersen, O. (1917). Negation in English and Other Languages. Copenhagen: Høst.
- Keizer, E. (2007). The lexical-grammatical dichotomy in Functional Discourse Grammar. Alfa Revista de Lingüística 51(2) 35-56.
- King, A., & Wedel, A. (2020). Greater early disambiguating information for less-probable words: The lexicon is shaped by incremental processing. Open Mind: Discoveries in Cognitive Science 4: 1-12.
- Klein, E. & I. Sag. (1985). Type-driven translation. *Linguistics and Philosophy* 8: 163-202.
- Koplenig, A., P. Meyer, S. Wolfer, & C. Müller-Spitzer. (2017). The statistical trade-off between word order and word structure Large-scale evidence for the principle of least effort. PLoS ONE 12(3): e0173614. doi:10.1371/journal. pone.0173614
- Kurumada, C. & S. Grimm. (2019). Predictability of meaning in grammatical encoding: Optional plural marking. Cognition 191: Article 103953.
- Kurumada, C., & Jaeger, T. F. (2015). Communicative efficiency in language production: Optional case-marking in Japanese. Journal of Memory and Language 83: 152–178.

REFERENCES (CONT.)

- Lehmann, C. (20153 [1982]). Thoughts on grammaticalization. Berlin: Language Science Press.
- Levy, R., & T. F. Jaeger. (2007). Speakers optimize information density through syntactic reduction. In B. Schlökopf, J. Platt, & T. Hoffman (eds.), Advances in Neural Information Processing Systems (NIPS). Cambridge, MA: MIT Press. 849-856.
- Martins, T. G., F. Simpson, F. Lindgren, & H. Rue. (2013). Bayesian computing with INLA: New features. *Comput. Stat. Data Anal.* 67: 68-83.
- Nichols, J. (1992). Linguistic diversity in time and space. Chicago, IL: University of Chicago Press.
- Montague, R. (1970). English as a formal language. In B. Visentini et al. (eds.), Linguaggi nella Societa et nella Technica. Milan: Edizioni di Communita. 188-221.
- Montague, R. (1973). The proper treatment of quantification in ordinary English. In K. J. J. Hintikka, J. M. E. Moravcsik, & P. Suppes (eds.), Approaches to Natural Language. Dordrecht: Reidel. 221–242.
- > Piantadosi, S. T., H. Tily, & E. Gibson. (2011). Word lengths are optimized for efficient communication. Proceedings of the National
- Academy of Sciences 108: 3526-3529.
- Roberts, S. G., & S. C. Levinson. (2017). Conversation, cognition and cultural evolution: A model of the cultural evolution of word order through pressures imposed from turn taking in conversation. Interaction Studies 18(3): 402–429.
- Siewierska, A. (1998). Variation in major constituent order: A global and a European perspective. In A. Siewierska (ed.), Constituent order in the languages of Europe: Empirical approaches to language typology. Berlin: Mouton de Gruyter. 475-551.
- Traugott, E. C. & E. König. (1991). The semantics-pragmatics of grammaticalization revisited. In E. C. Traugott & B. Heine (eds.), Approaches to grammaticalization. Volume I: Focus on theoretical and methodological issues. Amsterdam: Benjamins. 189-218.
- > Traugott, E. C. & G. Trousdale. (2013). Constructionalization and constructional changes. Oxford: Oxford University Press.
- > Zipf, G. K. (1929). Relative frequency as a determinant of phonetic change. Harvard Studies in Classical Philology 40: 1-95.
- > Zipf, G. K. (1949). Human behavior and the principle of least effort: An introduction to human ecology. Cambridge, MA: Addison-Wesley.
- Nichols, J. (1992). Linguistic diversity in time and space. Chicago, IL: University of Chicago Press.
- Pagel, M. & A. Meade. (2006). Bayesian analysis of correlated evolution of discrete characters by reversible-jump Markov chain Monte Carlo. Am. Nat. 167: 808-825.
- Piantadosi, S. T., H. Tily, & E. Gibson. (2011). Word lengths are optimized for efficient communication. Proceedings of the National Academy of Sciences 108: 3526–3529.

79

REFERENCES (CONT.)

- Roberts, S. G., & S. C. Levinson. (2017). Conversation, cognition and cultural evolution: A model of the cultural evolution of word order through pressures imposed from turn taking in conversation. Interaction Studies 18(3): 402–429.
- Siewierska, A. (1998). Variation in major constituent order: A global and a European perspective. In A. Siewierska (ed.), Constituent order in the languages of Europe: Empirical approaches to language typology. Berlin: Mouton de Gruyter. 475-551.
- Traugott, E. C. & E. König. (1991). The semantics-pragmatics of grammaticalization revisited. In E. C. Traugott & B. Heine (eds.), Approaches to grammaticalization. Volume I: Focus on theoretical and methodological issues. Amsterdam: Benjamins. 189-218.
- > Traugott, E. C. & G. Trousdale. (2013). Constructionalization and constructional changes. Oxford: Oxford University Press.
- > Zipf, G. K. (1929). Relative frequency as a determinant of phonetic change. Harvard Studies in Classical Philology 40: 1-95.
- Zipf, G. K. (1949). Human behavior and the principle of least effort: An introduction to human ecology. Cambridge, MA: Addison-Wesley.