

Sonia Cristofaro (University of Pavia)

Synchronic vs. diachronic explanations of typological universals: redefining the role of frequency

- (1) Three types of frequency effects on linguistic structure (Bybee 2001 and 2007, Mithun 2003, among others):
 - Phonetic reduction: high frequency items undergo phonetic reduction at a faster rate than low frequency ones.
 - Chunking: frequently occurring sequences of linguistic items come to be processed as a single chunk, leading to the loss of the identity of the component units.
 - Accessibility/autonomy: high frequency items become highly accessible and increasingly autonomous from related items.

These phenomena are plausibly all due to repetition, which leads to overlap and reduction of articulatory gestures, automatization of frequently occurring sequences of linguistic items, and strengthening of the memory representations for particular items.

- (2) In typology, frequency has also been invoked to account for a number of recurrent cross-linguistic asymmetries in the use of zero vs. overt marking for different grammatical categories (**structural markedness**: Greenberg 1966, Croft 2003):
 - Languages use zero marking for nominal, inanimate or indefinite direct objects ('I saw the boy/the tree/a man') and overt marking for pronominal, animate or definite ones ('I saw him/the boy/the man), but usually not the other way round.
 - Languages use zero marking for inalienable possession ('John's mother', 'John's arm') and overt marking for alienable possession ('John's books'), but usually not the other way round.
 - Languages use zero marking for singular and overt marking for plural, but usually not the other way round.
 - This has been accounted for by postulating an **economy** principle whereby overt marking is used for certain categories as opposed to others because these categories are less frequent, and hence more in need to be disambiguated from other categories. The categories that are zero marked, on the other hand, are frequent enough not to be in need of disambiguation, so they can be left unmarked (Greenberg 1966, Nichols 1988, Comrie 1989, Koptjevskaja-Tamm 1997, Dahl and Koptjevskaja-Tamm 1998 and 2001, Dixon 1994, Croft 2003, Haspelmath 2006 and 2008, among others).

- (3) The economy principle differs in several ways from traditional assumptions about phonetic reduction, chunking and accessibility/autonomy:
- Contrary to phonetic reduction, chunking and accessibility/ autonomy, the use of zero vs. overt marking is not a mechanical result of repetition and automatization. Rather, speakers selectively chose zero vs. overt marking based on some evaluation of the possible communicative effects of the relative frequency of particular categories (in terms of relative ambiguity of the category).
 - Phonetic reduction and chunking involve the loss of the original components of an expression. Economy, on the other hand, can in principle be implemented either through the loss of the markers for a more frequent category in a situation where all of the relevant categories are originally overtly marked, or through the development of overt markers for a less frequent category in a situation where all of the relevant category are originally zero marked.

But the economy principle has been proposed based on the synchronic cross-linguistic distribution of zero and overt marking for different categories, not the diachronic processes that actually give rise to this distribution in individual languages. Do these processes actually support the economy principle?

- Not really, in that in many cases the use of zero vs. overt marking for different categories originates through a variety of processes not obviously related to the relative frequency of these categories.
 - Also, these processes are rather different in nature and not obviously amenable to a unified explanation (**convergent evolution**: different developmental pathways from different sources give superficially similar results (Blevins 2004)).
- (4) Overt markers for less frequent categories when all of the relevant categories are originally zero marked:
- As is typically the case with overt markers, these often arise from the reinterpretation of elements originally used for other functions (through grammaticalization or other processes of form-meaning redistribution within complex expressions).
 - For example, in languages where all types of direct objects are zero marked, overt markers restricted to pronominal, animate, or definite direct objects can arise from topic markers (König 2008, Iemmolo 2010, among others; (5)).
 - In languages where both alienable and inalienable possession are zero marked, overt markers restricted to alienable possession can arise from locative expressions, e.g. ‘at’, ‘at X’s place, home’ and the like ((6), (7)).

- In languages where both singular and plural are zero marked, overt plural markers can arise from distributives ('house(s) here and there': (8)), or partitive expressions of the type 'many/ all of us' and the like, in which the quantifier is dropped and the plural meaning associated with it is transferred to a co-occurring element, for example a genitive case inflection originally indicating partitivity ((9)), or a verbal form ((10), where the plural marker was originally a participial form of the verb 'to be' used in expressions such as 'both of them', literally 'being two').

Kanuri (Nilo-Saharan)

- (5) (a) *Músa shí-ga cúro*
Musa 3SG-OBJ saw
'Musa saw him' (Cyffer 1998: 52)
- (b) *Káno-ro leji-ya ráwanzó súr-in*
kano-to go.3SG-DEP.FUT uncle see-IMPF
'When she goes to Kano, she will see her uncle' (Cyffer 1998: 70)
- (c) *wú-ga*
1SG-as.for
'As for me' (Cyffer 1998: 52)

Ngiti (Nilo-Saharan)

- (6) (a) *ma m-ìngyè àba bhà idzalí-nga*
1SG SC-be.i.the.habit.PFPR father POSS courtyard-NOMLZR
'I normally stay at the courtyard of my father' (Kutsch Lojenga 1994: 322)
- (b) *otsú-du*
hand-1SG.INAL.POSS
'my hand' (Kutsch Lojenga 1994: 202)
- (c) **bhà:**
at.home
'at home' (Kutsch Lojenga 1994: 154)

Manding (Niger-Congo)

- (7) (a) *buseo la sùboo*
butcher POSS meat
'the butcher's meat' (Creissels 2001: 441)
- (b) *nìnsoo sùboo*
cow flesh
'the cow's flesh' (Creissels 2001: 441)
- (c) *Mùrú ye galà la*
knife.DEF be platform.DEF on
'The knife is on the platform.' (Creissels 2009: 129; Creissels (2001:

453) reconstructs the original meaning of *la* as ‘mouth, border’, and argues that the possessive construction ‘X *la* Y’ was originally ‘the Y at X’s’)

Southern Paiute (Uto-Aztecan)

- (8) (a) *qa’ni* / **qaŋqa’ni**
house / house.DISTR
‘house, houses’ (Sapir 1930-1: 258)
- (b) *piŋwa-* / **pivi’ŋwa.mi**
wife / wife.DISTR.their
‘wife / their (vis.) wives’ (Sapir 1930-1: 257)

Bengali (Indo-European)

- (9) (a) *chēlē-rā*
child-GEN
‘children’ (15th century: Chatterji 1926: 736)
- (b) *āmhā-rā sãbã*
we-GEN all
‘all of us’ (14th century: Chatterji 1926: 735)

Assamese (Indo-European)

- (10) (a) *chātar-hāt*
student-PL
‘Students’ (Modern Assamese: Kakati 1962: 295)
- (b) *dui-hanta*
two-be.PTCPL
‘Both of them’ (Early Assamese: Kakati 1962: 283)

(11) Processes such as those in (4) have long been described in classical historical linguistics and grammaticalization studies, and they are generally regarded as a result of context-driven mechanisms:

- In some cases, for example, some meaning component is transferred from one formal component of a complex expression to another:
 - Topic markers presumably become direct object markers when they are used with topicalized direct objects, and take on the direct object meaning of the latter.
 - Partitive case endings and verbal forms become plural markers as they take on a plural meaning originally associated with a co-occurring quantifier.
- In other cases, some element becomes associated with a meaning component that can naturally be inferred from the context:

- Possession can be inferred in many contexts involving locative expressions (e.g., ‘the courtyard’s in my father’s house’ > ‘my father’s courtyard’, ‘the meat at the butcher’s’ > ‘the butcher’s meat’: Claudi and Heine 1986, Heine, Claudi, and Hünemeyer 1991: chap. 6), so these expressions can easily develop a possessive meaning.
 - Distributives can develop a plural meaning because, when applied to individuated items, they always involve the notion of plurality (Mithun 1999: 90; similar observations apply to many other possible sources for plural markers, e.g. expressions meaning ‘several’, ‘all’, ‘people’, see Cristofaro 2013 and 2014 for exemplification and discussion).
- In all of these cases, there is no obvious evidence that the development of the relevant overt markers is motivated by the lower frequency of the categories that they encode and the need to disambiguate these categories. Rather, this development is a result of contextually dependent associations that speakers establish between those categories and highly specific source elements (metonymization: Heine, Claudi, and Hünemeyer 1991, Bybee, Perkins, and Pagliuca 1994, Heine 2003, Traugott and Dasher 2005, among others).
 - The categories not involved in the process (e.g. singular as opposed to plural) retain zero marking, which was the strategy originally used for all categories, so in this case too there is no evidence that the use of zero marking for these categories is motivated by their higher frequency and their lower need for disambiguation.
- (12) ‘Distributives usually imply plurality [...] At a certain point the implied plurality is reanalyzed as the central meaning of the form.’ (Mithun 1999: 90)
- (13) In fact, specific restrictions in the distribution of overt markers can be directly related to the properties of the source construction, independently of the frequency of the relevant categories:
- Topics are usually pronominal, animate, and definite, so it should be expected that direct object markers derived from topic markers will be restricted to pronominal/animate/definite direct objects, at least initially. This is independent of the greater vs. lower need to disambiguate different types of direct objects.
 - Locative expressions are not usually used to refer to inalienably possessed items (? ‘The mother in John’s house’, ? ‘The arm at John’s’), so it should be expected that the resulting possessive markers will be (initially) restricted in the same way, independently of the greater vs. lower need to disambiguate different possession types. In some languages, possessive constructions derived from locative expressions are or are not used to encode alienable possession

depending on whether or not the specific relationship between possessor and possessee is compatible with a locative meaning ((14)).

Manding (Niger-Congo)

(14) (a) *à màario/ bàtufaa/ kàammoo*
3SG master patron teacher

‘his master/ patron (of a griot)/ teacher (in a Koranic school)’
(Creissels 2001: 446)

(b) *à la jòŋo/ jàloo/ kàrandiŋo*
3SG POSS slave griot pupil

‘his slave/ griot/ pupil (in a Koranic school)’ (Creissels 2001: 446)

(15) Depending on the properties of the source construction, individual processes can give rise to overt marking for different categories, independently of the frequency of these categories:

- Direct object markers derived from the reinterpretation of constructions other than topic markers and not restricted to pronominal, animate or definite elements have no such restrictions either (e.g. markers derived from ‘take’ verbs in constructions of the type ‘take X and Verb (X)’ as these are reinterpreted as ‘OBJ X Verb’: (16); markers derived from possessor markers in nominalized constructions of the type ‘the making of X’ as these are reinterpreted ‘make X’: (17)).
- Partitive constructions with quantifiers referring to singular items (‘one of them’ and the like) give rise to singular, rather than plural markers ((18): plural is zero marked in the language).
- When the source construction is compatible with both singular and plural, it gives rise to overt markers for both. Sometimes, for example, number markers are actually gender markers that developed from demonstratives and personal pronouns with distinct singular and plural forms. In such cases, there are overt markers for both singular and plural (table (1)).

Ga (Niger-Congo)

(16) *è kè nù wò tó lè mlĩ*
she OBJ water put bottle the inside

‘She put water in the bottle.’ (Lord 1993: 119; cf. Yoruba *ká, kò*, Fon *kple*, Ewe *kè*, ‘pick, take, gather, collect’: Lord 1993: 120)

Wayana (Carib)

(17) *i-pakoro-n iri-Ø pək wai*
1-house-POSS/OBJ make-NOMLZR occ.with 1.be

‘I’m (occupied with) making my house (originally, ‘my house’s making’)
(Gildea 1998: 201)

Imonda (Border)

- (18) (a) *agō-ianèi-m ainam fa-i-kōhō*
 women-NONPL-GL quickly CL-LNK-go
 ‘He grabbed the woman’ (Seiler 1985: 194)
- (b) *mag-m ad-ianèi-m*
 one-GL boys-NONPL-GL
 ‘To one of the boys’ (Seiler 1985: 219)
- (c) *po me-ianèi*
 water hole-SRC
 ‘from underneath the water’ (Seiler 1985: 73)

		SG	PL	
Nouns	M	/ǒá- mà	/ǒá-// u‘a	‘boy’
	F	/ǒá- hè	/ǒá- djì	‘girl’
	C	/ǒá-(‘à), /ǒá- djì	ǒá- nà	‘child’
Pronouns	M	xà- má , á- mà , i- mà	xà-// uá , á-// uá , í-// uá	‘he’
	F	xà- hè , á- hè , i- hè	xà- djì , á- djì , í- djì	‘she’
	C	(xa-’à)	xà- nà , á- nà , í- nà	‘it’

Table 1: Gender/number markers and third person pronouns in Kxoe (Khoisan: Heine 1982: 211)

- (19) Zero marking for a more frequent category when all of the relevant categories are originally overtly marked:

- Sometimes, this is the result of regular phonetic changes leading to the elimination of the marker originally used for the more frequent category.
- In English, for example, the current configuration with zero marked singulars and *-s* marked plurals (traditionally regarded as a textbook case of economy) resulted from a series of phonetic changes that led to the elimination of all inflectional endings except genitive singular *-s* and plural *-es* (Mossé 1949; table 2).
- As phonetic changes are arguably independent of the categories encoded by the affected forms, such cases provide no evidence that the resulting patterns are related to the relative frequency of the relevant categories. In fact, cross-linguistically, such changes can also lead to patterns with overt marking for the more frequent category and zero marking for the less frequent one (e.g. overtly marked singulars and zero marked plurals: (20), (21)).

		Class I	Class II	Class III
SG	NA	-	-e	-e
	G	-(e)s	-es	-e
	D	-e	-e	-e
PL		-(e)s	-en GEN	-en(e)

Table 2: Nominal declension in Middle English (Mossé 1949: 65)

- (20) Sinhala (Indo-European): some inanimate nouns have overtly marked singulars and zero marked plurals (e.g. *pot-a/ pot* ‘book-SG/ book.PL’). This was a result of phonetic changes leading to the loss of the plural ending of a specific inflectional class in the ancestor language (Nitz and Nordhoff 2010: 250-6).
- (21) Nchanti (Niger-Congo): Nouns in classes 3/4 have overt marking in the singular and zero marking in the plural, e.g. *k^wəŋ/ kəŋ* ‘firewood.SG/ firewood.PL, *k^wēē/ kēē* ‘moon.SG/ moon.PL’. Originally, both singular and plural were marked overtly through the two prefixes **u-* and **i-* respectively. As these were eliminated, the singular prefix led to the labialization of the initial consonant of the stem, while the plural prefix left no trace (Hombert 1980).
- (22) Some cases where the development of zero marking can be directly related to the frequency of the relevant categories:
- Sometimes, forms with overt marking for a more frequent category (third person, absolutive) are reanalysed as a single whole, so that the marker does not function as such any more, and the category becomes zero marked. The form is then used as a basis to rebuild other, related forms through the addition of overt markers for the relevant categories (Koch 1995, Bybee 2007, among others: table 3).
 - These processes are presumably related to frequency, but not in terms of economy. The reanalysis of individual forms as a single whole is an instance of chunking, so it’s plausibly due to repetition and automatization. The fact that the form is selected as a basis for other forms is an effect of its relatively high degree of autonomy and accessibility, again due to repetition (Bybee 2007).
- (23) Concluding remarks:
- Diachronic evidence suggests that the distribution of zero vs. overt marking across different grammatical categories depends on what processes and source constructions give rise to zero vs. overt marking, rather than the relative frequency of these categories.

	'this' I	'this' II
ABS	nhe-nhe	nhenhe
ERG	nhe-le	nhenhe-le
DAT	nhe-ke	nhenhe-ke

Table 3: The evolution of the paradigm for 'this' in Alyawarre (Koch 1995: 39)

- Also, synchronically based explanations postulate a single principle (economy) to account for all of the cases where a more frequent category is zero marked and a less frequent one is overtly marked, but
 - These cases can be a result of very different processes, such as phonetic changes, meaning transfer from one component of a complex expression to another, grammaticalization, or chunking. From the diachronic point of view, it is difficult to relate all of these processes to a single overarching principle.
 - Many of these processes lead to different, sometimes opposite distributional patterns for zero vs. overt marking depending on the source construction.
- A full understanding of the distribution of zero vs. overt marking, then, requires diachronic data about the particularized, often context-dependent processes that give rise to the relevant patterns (for example, what source constructions can give rise to specific overt markers, in what contexts, through what mechanisms), rather than data about the frequency of the categories being expressed.
- In general, this supports two points occasionally raised by some typologists (Bybee 1988, 2006 and 2008; Aristar 1991). First, typological explanations should refer to the processes and constructions that actually give rise to particular patterns diachronically, rather than on the resulting patterns in themselves. Second, individual distributions may be the result of several distinct processes, not necessarily amenable to a unified explanation. This line of research has not been systematically explored in typology, but has a parallel, for example, in Evolutionary Phonology (Blevins 2004).

Abbreviations

	GL	goal	OBJ	object	
C	common	IMPF	imperfect	PFPR	perfective present
CL	classifier	INAL	inalienable	POSS	possessive
DEF	definite	LNK	link	POSS	possessor
DEP.FUT	dependent future	M	masculine	PTCPL	participle
DISTR	distributive	NOMLZ	nominalizer	SG	singular
F	feminine	NOMLZR	nominalizer	SRC	source
GEN	genitive	NONPL	non-plural		

References

- Aristar, A. R. (1991). On diachronic sources and synchronic patterns: an investigation into the origin of linguistic universals. *Language* 67, 1–33.
- Blevins, J. (2004). *Evolutionary phonology: the emergence of sound patterns*. Cambridge: Cambridge University Press.
- Bybee, J. (1988). The diachronic dimension in explanation. In J. A. Hawkins (Ed.), *Explaining language universals*, pp. 350–79. Oxford: Basil Blackwell.
- Bybee, J. (2001). *Phonology and Language Use*. Cambridge: Cambridge University Press.
- Bybee, J. (2006). Language change and universals. In R. Mairal and J. Gil (Eds.), *Linguistic Universals*, pp. 179–94. Cambridge: Cambridge University Press.
- Bybee, J. (2007). *Frequency of Use and the Organization of Language*. Oxford: Oxford University Press.
- Bybee, J. (2008). Formal Universals as Emergent Phenomena: The Origins of Structure Preservation. In J. Good (Ed.), *Linguistic Universals and Language Change*, pp. 108–21. Oxford: Oxford University Press.
- Bybee, J., R. Perkins, and W. Pagliuca (1994). *The evolution of grammar*. Chicago and London: The University of Chicago Press.
- Chatterji, S. K. (1926). *The Origin and Development of the Bengali Language*. Calcutta: Calcutta University Press.
- Claudi, U. and B. Heine (1986). On the metaphorical base of grammar. *Studies in Language* 10, 297–335.
- Comrie, B. (1989). *Language universals and linguistic typology. 2nd edition*. Oxford: Basil Blackwell.
- Creissels, D. (2001). Catégorisation et grammaticalisation: la relation génitive en mandingue. In R. Nicolaï (Ed.), *Leçons d’Afrique (hommage à Gabriel Manessy)*, pp. 433–54. Leuven - Paris: Peeters.
- Creissels, D. (2009). *Le malinké de Kita*. Köln: Rüdiger Koppe.

- Cristofaro, S. (2013). The referential hierarchy: Reviewing the evidence in diachronic perspective. In D. Bakker and M. Haspelmath (Eds.), *Languages across Boundaries: Studies in the Memory of Anna Siewierska*, pp. 69–93. Berlin and New York: Mouton de Gruyter.
- Cristofaro, S. (2014). Competing motivations and diachrony: what evidence for what motivations? In B. MacWhinney, A. Malchukov, and E. Moravcsik (Eds.), *Competing motivations in grammar and usage*, pp. 282–98. Oxford: Oxford University Press.
- Croft, W. (2003). *Typology and universals. 2nd edition*. Cambridge: Cambridge University Press.
- Cyffer, N. (1998). *A Sketch of Kanuri*. Köln: Rüdiger Köppe.
- Dahl, Ö. and M. Koptjevskaja-Tamm (1998). Alienability splits and the grammaticalization of possessive constructions. In *Papers from the XVIth Scandinavian Conference of Linguistics*, Turku. University of Turku and Åbo Akademi.
- Dahl, Ö. and M. Koptjevskaja-Tamm (2001). Kinship in grammar. In I. Baron, M. Herslund, and F. Sørensen (Eds.), *Dimensions of Possession*. Amsterdam and Philadelphia: John Benjamins.
- Dixon, R. M. W. (1994). *Ergativity*. Cambridge: Cambridge University Press.
- Gildea, S. (1998). *On reconstructing grammar : Comparative Cariban morphosyntax*. Oxford: Oxford University Press.
- Greenberg, J. H. (1966). *Language universals, with particular reference to feature hierarchies*. The Hague: Mouton.
- Haspelmath, M. (2006). Against markedness (and what to replace it with). *Journal of Linguistics* 42, 25–70.
- Haspelmath, M. (2008). Creating economical morphosyntactic patterns in language change. In J. Good (Ed.), *Linguistic Universals and Language Change*, pp. 185–214. Oxford: Oxford University Press.
- Heine, B. (1982). African Noun Class Systems. In H. Seiler and C. Lehmann (Eds.), *Apprehension : das sprachliche Erfassen von Gegenständen*, pp. 189–216. Tübingen: Narr.
- Heine, B. (2003). Grammaticalization. In B. D. Joseph and R. D. Janda (Eds.), *The Handbook of Historical Linguistics*, pp. 576–601. Oxford: Blackwell.
- Heine, B., U. Claudi, and F. Hünnemeyer (1991). *Grammaticalization*. Chicago: University of Chicago Press.
- Hombert, J.-M. (1980). Noun Classes of the Beoid Languages. *Southern California Occasional Papers in Linguistics* 8, 83–98.
- Iemmolo, G. (2010). Topicality and differential object marking: evidence from Romance and beyond. *Studies in Language* 34.2, 239–72.
- Kakati, B. (1962). *Assamese , its formation and development. 2nd edition*. Lawyer's Book Stall: Gauhati.

- Koch, H. (1995). The creation of morphological zeroes. Volume 1994 of *Yearbook of Morphology*, pp. 31–71. Dordrecht: Kluwer.
- König, C. (2008). *Case in Africa*. Oxford: Oxford University Press.
- Koptjevskaja-Tamm, M. (1997). Possessive noun phrases in Maltese: Alienability, iconicity and grammaticalization. *Rivista di Linguistica* 8, 245–74.
- Kutsch Lojenga, C. (1994). *Ngiti : a central-Sudanic language of Zaire*. Köln: Rüdiger Köppe.
- Lord, C. (1993). *Historical change in serial verb constructions*. Amsterdam and Philadelphia: John Benjamins.
- Mithun, M. (1999). *The Languages of Native North America*. Cambridge: Cambridge University Press.
- Mithun, M. (2003). Functional Perspectives on Syntactic Change. In R. D. Janda and B. D. Joseph (Eds.), *The handbook of historical linguistics*, pp. 552–72. Oxford: Blackwell.
- Mossé, F. (1949). *Manuel de l'anglais du Moyen âge des origines au XIVme siècle. II. Moyen-Anglais. Tome premier: Grammaire et textes*. Paris: Aubier.
- Nichols, J. (1988). On alienable and inalienable possession. In W. Shipley (Ed.), *In Honor of Mary Haas*, pp. 557–609. Berlin: Mouton de Gruyter.
- Nitz, E. and S. Nordhoff (2010). Subtractive Plural Morphology in Sinhala. In J. Wohlgemuth and M. Cysouw (Eds.), *Rara & Rarissima: Collecting and interpreting unusual characteristics of human languages*, pp. 247–66. Berlin and New York: Mouton de Gruyter.
- Sapir, E. (1930–1931). *The Southern Paiute Language*. Boston: American Academy of Arts and Sciences.
- Seiler, W. (1985). *Imonda, a Papuan language*. Pacific Linguistics. Series B-93. Canberra: The Australian National University.
- Traugott, E. C. and R. B. Dasher (2005). *Regularity in Semantic Change*. Cambridge: Cambridge University Press.