

# A Frequentist Explanation of Some Universals of Reflexive Marking\*

Martin Haspelmath

Max Planck Institute for Evolutionary Anthropology

*This paper identifies a number of empirically observable universals of reflexive marking that concern the existence of a special reflexive pronoun and the length of the marker that is used in reflexive constructions, in various different positions of the nonreflexive or reflexive pronoun. Most of the proposed universals have been mentioned earlier in the literature, but they have not been very prominent because the literature on binding has focused on language-specific generalizations rather than identifying readily testable cross-linguistic generalizations. I argue that all of these universals have their basis in a frequency asymmetry: Under different circumstances, the likelihood of an anaphoric pronoun being coreferential with the subject can be quite different, and this is argued to be the motivation for the universal patterns of form.*

## 1. Some explananda

In this paper, I propose frequency-based explanations of a number of universal contrasts in the form of reflexive marking, and I compare them with widely assumed generative explanations. The methodology of my approach is rather different from the well-known generative approach, so I make its methodological foundations explicit in §2, before advancing and discussing seven universals of reflexive marking in §3-7. In §8 I briefly address further methodological implications and draw conclusions.

To get a first idea of the kinds of phenomena that I will deal with in this paper, let us look at some potentially relevant contrasts of reflexive marking. First, some languages such as Russian have a contrast between **introverted** (usually self-directed) and **extroverted** (usually other-directed) verbs (Haiman 1983:803), whereas other languages such as German lack this contrast. This is seen in (1)-(2).

- (1) Russian
- a. Vanja moet-sja. (#moet sebja)  
'Vanja washes himself.' ('washes himself')
  - b. Vanja nenavidit sebja (\*nenavidit-sja)  
'Vanja hates himself.'
- (2) German
- a. Gertrud wäscht sich.  
'Gertrud washes (herself).'
  - b. Gertrud hasst sich.  
'Gertrud hates herself.'

---

\* Earlier versions of this paper were presented at the Workshop on Reciprocity and Reflexivity (FU Berlin, October 2004), at the SLE conference in Valencia (September 2005), at summer schools at MIT (2005) and in Campobasso (2007), and at the Max Planck Institute for Evolutionary Anthropology. I am grateful for useful comments to the audiences at these occasions, to Peter Cole, Gabriella Hermon and Kim Schulte, and to several anonymous reviewers.

In Russian, the reflexive suffix *-sja* can only be used with introverted verbs like ‘wash’ (see 1a), whereas extroverted verbs like ‘hate’ must use the free reflexive pronoun *sebjja* (see 1b).<sup>1</sup> (The free reflexive pronoun is also possible in (1a), but in general only in a contrastive context, indicated by “#.”)

Another contrast made by some languages is that between **direct object** and **adnominal possessor** of direct object, as in English (3a-b), where the direct-object pronoun (*himself*) must be marked as reflexive when coreference with the subject is intended, but the adnominal possessor cannot be marked as possessive (*\*himself’s*). In other languages such as Lezgian (4a-b), both constructions require a reflexive pronoun.

- (3) English
- a. Bob<sub>1</sub> admires himself<sub>1</sub>. (\*Bob<sub>1</sub> admires him<sub>1</sub>)
- b. Bob<sub>1</sub> admires his<sub>1</sub> boss. (\*Bob admires himself’s boss.)
- (4) Lezgian (Haspelmath 1993:408-414)
- a. Ali-diz wič akuna. (\*Alidiz<sub>1</sub> am<sub>1</sub> akuna)  
 Ali-DAT self saw him  
 ‘Ali saw himself.’
- b. Ali-diz wič-in ruš akuna. (\*Alidiz<sub>1</sub> adan<sub>1</sub> ruš akuna)  
 Ali-DAT self-GEN girl saw  
 ‘Ali<sub>1</sub> saw his<sub>1</sub> daughter.’

In (4a-b), the nonreflexive pronoun *am/adan* clearly indicates a noncoreferential reading.

Third, while many languages have a clear contrast between **disjoint reference** and **coreference** as it is well known from English (5a-b), other languages such as the Oceanic language Loniu (spoken in Papua New Guinea) do not show such a contrast (cf. 6a-b).

- (5) English
- a. Bob<sub>1</sub> saw him<sub>2</sub>.
- b. Bob<sub>1</sub> saw himself<sub>1</sub>. (\*Bob<sub>1</sub> saw him<sub>1</sub>)
- (6) Loniu (Hamel 1994:54)
- a. Suʔu<sub>1</sub> čaʔiti suʔu<sub>2</sub>.  
 they.DU cut they.DU  
 ‘They cut them.’
- b. Suʔu<sub>1</sub> čaʔiti suʔu<sub>1</sub>.  
 they.DU cut they.DU  
 ‘They cut themselves/each other.’

Loniu has no reflexive pronouns and constructions with anaphoric pronouns in object position like (6) are vague with respect to the disjoint reference/ coreference distinction.

Finally, in some languages reflexive pronouns can only be used when there is **full coreference**. This is the case in English, where (7b) with **partial coreference** of the reflexive

<sup>1</sup> König & Siemund (2000a) avoid Haiman’s terms *introverted/extroverted* (in favor of *non-other-directed/other-directed*) because they feel that the semantic categories are too different from the relevant notions in (popular) psychology. My sense is that there is no real danger of misunderstanding, so I retain Haiman’s original terminology.

pronoun is impossible (the intended reading is one where Maria criticizes a group of people that she is a member of). In Hausa, by contrast, the ordinary reflexive pronoun can be used in such cases of partial coreference (see 8b).

- (7) English  
 a. Maria<sub>1</sub> criticized herself<sub>1</sub>.  
 b. \*Maria<sub>1</sub> criticized themselves<sub>1+x</sub>. ('herself and the others')
- (8) Hausa (Newman 2000:524)  
 a. Laadi<sub>1</sub>      taa      soòki      káàn-tà<sub>1</sub>.  
    Ladi            3SG      criticize    self-3PL  
    'Ladi criticized herself.'  
 b. Laadi<sub>1</sub>      taa      soòki      káàn-sù<sub>1+x</sub>.  
    Ladi            3SG      criticize    self-3SG.F  
    '(lit.) Ladi criticized themselves.'

Anyone who is interested in explaining the phenomena of reflexive marking in human languages has to confront the problem that none of the contrasts illustrated in (1), (3), (5) and (7) is a necessary feature of languages, as the examples in (2), (4), (6) and (8) show. So how is explanation possible in this domain? After all, "explaining something" basically means showing that it necessarily follows from something more general.

The next section will compare two different modes of approaching this general problem. In the later sections, we will see that a frequency-based explanation is available for the contrasts in (1)-(2), (3)-(4), and (5)-(6), but not yet for the contrast in (7)-(8). That usage frequency is an important ingredient in an explanatory account of reflexive behavior is apparently a very new idea. Only Ariel (2008: ch. 6) adopts a very similar usage-based approach (her and my accounts were originally developed independently from each other).

To conclude this introductory section, a note on terminology: I use the term *phoric pronoun* (or simply *phoric*) as a cover-term for discourse-referring anaphoric/cataphoric pronouns ("personal pronouns," "anaphoric demonstratives") and strictly intrasentential pronouns ("reflexive pronouns," "anaphors"). The term *reflexive* is used for a form (a phoric pronoun or a verbal marker) that expresses coreference of a notional participant with the subject (or one of the subjects) of the sentence (among other functions; such forms commonly have other functions as well).

## 2. Some methodology: Two modes of explanation

The two modes of explanation that I will briefly contrast here are the **generative** mode and the **functional-frequentist** mode. My own frequency-based explanation instantiates the latter mode.

In generative syntax, explanation in the face of cross-linguistic variation is generally achieved by (i) observing a language-particular generalization, (ii) making claims about the mental grammar underlying it, (iii) deriving much of the language-particular grammar from a restrictive model of Universal Grammar (UG), (iv) observing new language-particular facts inconsistent with the model of UG, (v) proposing a revised model of UG that allows for all known grammars (but not more), and (vi) repeating steps (iv) and (v) over and over (hoping that the model of UG will remain restrictive). The idea is that to the extent that the final model of UG excludes certain logically possible language types, the non-existence of these types is explained (cf. Haspelmath to appear for further discussion).

In the functional-frequentist (or usage-based<sup>2</sup>) approach, by contrast, explanation is achieved by (i) examining the phenomenological grammars of a wide variety of languages (ii) formulating inductive cross-linguistic generalizations (= empirical universals), and (iii) proposing functional explanations of the observed universals, i.e. explanations that derive the universals from more general aspects of language use. No claims about language-particular mental grammars or Universal Grammar are needed (see Haspelmath 2004 for a detailed defense of this claim). The idea is that to the extent that a language-particular pattern instantiates an explained universal, it has been explained (in a weak sense).<sup>3</sup>

In the following sections, I will formulate seven universals of reflexive marking, and I will propose frequency-based explanations for them. The proposed universals have not really been substantiated by a world-wide study, but most of them have long been discussed in the literature, so any serious counterevidence would probably have come to light by now. The statistical data adduced here are not yet sufficient to show conclusively that the frequency trends are indeed as universal as I claim they are. Ideally one would like to have corpus data from a wide range of diverse languages (representing spontaneous everyday speech), but such data are not currently available, or at least not readily accessible. So I follow Postal (1970) in spirit and make some strong universal claims on the basis of limited evidence, hoping that others will thereby be challenged to look for confirming evidence or counterevidence (depending on whether they find my overall story attractive or not).

It should be noted that I do not claim that all universals of reflexive marking can be explained in frequentist terms. Some syntactic-semantic asymmetries are not a matter of more or less, but of yes or no. For example, 3rd person phoric pronouns offer multiple referential choices, whereas the reference of 1st person pronoun is always clear. This has been said to explain the well-known universal that if a language has a first person reflexive pronoun, it also has a third person reflexive pronoun (Faltz 1985: 43, 120; Comrie 1989: 6-7, 28; 1999:337).<sup>4</sup>

Nor do I claim that the universals of reflexive marking discussed here are the most important ones or are in other ways representative. This would be totally inappropriate, because in this paper I do not consider universals of the syntactic-semantic relation between the antecedent and the reflexive (“o-/a-/c-command”), an area of much research on reflexives. At the moment I am agnostic about how these universals are best formulated and how they should be explained,<sup>5</sup> and I concentrate on the more tractable phenomena concerning the form of the reflexive marker.

---

<sup>2</sup> The term *usage-based* is currently more common in linguistics than *frequentist*, but since there is also usage-based work that does not appeal to frequency information directly (e.g. Langacker 2000), I use the more explicit term *frequentist* here (cf. also Mattausch 2007).

<sup>3</sup> See Vennemann (1983) on weak vs. strong explanation in linguistics. In this conception, only universals can be explained in a strong sense, because only universals are properties of human language in general.

<sup>4</sup> This explanation of Faltz’s universal has recently been challenged by Newmeyer (2003:694-5), who points out that there is an alternative functional explanation available: It could be that the greater frequency of third person pronouns is the sole explanation of their greater formal differentiation, because in general more frequently used concepts are more likely to be lexicalized than less frequently used concepts. If Newmeyer is right, then this asymmetry, too, receives a frequency-based explanation.

<sup>5</sup> Although there seems to be some trends, as far as I can determine, the literature does not contain clear testable universal claims about antecedent-phoric relations. A widespread view in the generative literature is that a version of Chomsky’s (1981) binding conditions is universal and part of the innate Universal Grammar. However, the usual versions of the binding conditions do not provide independent, universally applicable definitions of the categories that figure in them (“anaphor,” “pronominal”), so that the binding conditions cannot be empirically tested with cross-linguistic data. If confronted with a phoric expression that behaves unexpectedly, one could always claim that it is neither an anaphor nor a pronominal and hence simply does not fall under the binding conditions (as is done, for example, in Cole & Hermon 2005 with respect to Malay

### 3. Introverted vs. extroverted actions

The first universal to be mentioned here has been known since Faltz (1985) (first circulated in 1977), was prominently discussed by Haiman (1983:801-08), and more recently by König & Siemund (2000a), König & Vezzosi (2004), and Smith (2004). It concerns the form of reflexive marking in simple transitive constructions.

#### (9) Universal 1

In all languages, the reflexive-marking forms employed with extroverted verbs are at least as long (or “heavy”) as the reflexive-marking forms employed with introverted verbs.

Typical extroverted verbs are transitive verbs like ‘kill’, ‘hate’, ‘criticize’, ‘see’, ‘attack’, and typical introverted verbs are verbs like ‘wash’, ‘shave’, ‘dress’, ‘defend’. Some examples of languages showing an introversion/extroversion contrast are given in Table 1. Most of these have been so widely discussed in the literature on reflexives (e.g. Faltz 1985, Haiman 1983, Geniušienė 1987, König & Siemund 2000a) that no further references are necessary here.<sup>6</sup>

	EXTROVERTED		INTROVERTED	
English	hate <b>onself</b>		shave <b>Ø</b>	
Russian	nenavidet’ <b>sebja</b>	‘hate oneself’	myt’- <b>sja</b>	‘wash’
Hungarian	utálja <b>mag-á-t</b>	‘hates herself’	borotvál- <b>koz-</b>	‘shave’
Greek	aghapái <b>ton eaftó tu</b>	‘loves himself’	dín- <b>ete</b>	‘dresses’
Turkish	<b>kendini</b> sev-iyor	‘loves himself’	yıka- <b>n-iyor</b>	‘washes’
Dutch	haat <b>zichzelf</b>	‘hates herself’	wast <b>zich</b>	‘washes’
Frisian	hearde <b>himsels</b>	‘heard himself’	wasket <b>him</b>	‘washed’
Romanian	se vede pe <b>sine însuși</b>	‘sees himself’	<b>se</b> spală	‘washes’
Jamul Tiipay	<b>naynaach mat</b> -aaxway	‘killed himself’	<b>mat</b> -sxwan	‘scratch (oneself)’

Table 1: Extroverted and introverted reflexive forms in some languages

The formal types of contrasts between extroverted and introverted constructions are quite diverse: English shows a contrast between an overt pronoun and nothing, Russian, Hungarian and Turkish have a contrast between a pronoun and a verbal affix, Dutch has a contrast between a longer and a shorter reflexive pronoun, and Frisian has a contrast between a longer reflexive and a shorter nonreflexive pronoun. Still, all these cases fall under Universal 1.<sup>7</sup> Languages lacking an introversion/extroversion contrast (like German) do not contradict Universal 1, although they provide no evidence for it.

A generative explanation of Universal 1 has not to my knowledge been proposed so far. The introverted/extroverted contrast has been discussed by Everaert (1986) and Reinhart & Reuland (1993:666) for Dutch and Frisian, but a purely stipulative account has been offered by these authors; in Reinhart & Reuland’s terms, introverted verbs have two lexical entries, one of which is “lexically reflexive.” This would allow a hypothetical but unattested language in which extroverted verbs like ‘hate’ and ‘see’ are lexically reflexive and hence receive short reflexive marking, while introverted verbs like ‘wash’ and ‘dress’ are not lexically reflexive

*dirinya*). In effect this means that the binding conditions are merely claims about English, not about Universal Grammar.

<sup>6</sup> The Jamul Tiipay (Yuman; southern California) data are from Miller (2001:166-167), and the Romanian data are from Calude (2007:242, 246).

<sup>7</sup> “Long” in Universal 1 can be interpreted as referring to simple segmental length. Ultimately the relevant factor is probably articulatory effort, and this may not always correlate exactly with segmental length. But for the purposes of this paper, segmental length suffices as an approximation.

and hence require long reflexive marking. But it is precisely such languages that Universal 1 excludes.

The functional explanation has been stated clearly by Haiman (1983): It is the principle of economical coding of predictable information (“What is predictable receives less coding than what is not,” Haiman 1983:807). But why exactly is the reflexive interpretation of introverted verbs predictable? I claim that it is the relative frequency of reflexive use of a given verb. If a verb is rarely used reflexively, marking it as reflexive is more important than if a verb is often used reflexively.

The relevance of frequency has been implicit in the literature since Faltz’s groundbreaking work. Introverted verbs have been characterized as “verbs expressing *commonly* reflexive actions such as washing oneself” (Faltz 1985:8), as expressing “*normally* reflexive activities” (Faltz 1985:19), “actions which one *generally* performs upon one’s self” (Haiman 1983:803), “*stereotypically* reflexive actions” (Levinson 2000:329), or actions “*typically* or *conventionally*” directed at oneself (König & Siemund 2000a:60; emphasis added in all cases). But for some reason, linguists have generally been reluctant to define introverted verbs as those that occur with high frequency in reflexive use, and to correlate forms directly with frequencies. König & Siemund (2000a:60-61) talk about introversion/extroversion as a “semantic property” or as involving “world knowledge.”

But verb meaning does not seem to be the decisive factor: In a hypothetical culture where people are always shaved by others, a verb meaning ‘shave’ would not behave as an introverted verb, even if it were semantically fully identical to English *shave*. So is frequency in the world (or knowledge of that frequency, i.e. world knowledge) the crucial quantity? But how would world frequency get reflected in language structure?

Clearly, the mechanism for economic motivation of the sort discussed by Haiman (and Zipf before him, cf. Zipf 1935) is the grammaticalization of speakers’ tendencies in discourse. Speakers can afford to reduce expressions that hearers can predict they will hear, but they have to be fully explicit on expressions that surprise hearers because of their rarity. Structural Zipfian economy derives from **speech frequency**, not from world frequency. Often speech frequency correlates with and is due to world frequency (as presumably in the case of introverted/extroverted verbs), but in many other cases world frequency has no relation to speech frequency (for example, the word *oxygen molecule* is rarer than the word *house*, although houses are much rarer in the world; and plurals are rarer than singulars, although the world contains more groups than individuals; see also Ariel (2008) and Haspelmath (2008) for related discussion).

That speech frequency, not world frequency, is the immediately relevant factor is fortunate, because unlike world frequency, it can be measured rather easily, by performing frequency counts of representative text corpora. So is it true that introverted verbs occur “typically” or “normally” reflexively?

I did a very simple corpus search using the on-line version of the British National Corpus and found the figures in Table 2.

extroverted: <i>kill</i>	disjoint ('kill someone')	86	<b>(79%)</b>	(full NP object: 59)
	coreferential ('kill oneself')	5	<b>(5%)</b>	(pronoun object: 27)
	objectless ('be a killer')	18	<b>(17%)</b>	
introverted: <i>wash</i>	disjoint ('wash someone')	35	<b>(70%)</b>	(full NP object: 28)
	coreferential ('wash onself')	11	<b>(22%)</b>	(pronoun object: 7)
	objectless ('be a washer')	4	<b>(8%)</b>	

Table 2: Different transitive verbs with coreferential and disjoint objects (source: British National Corpus)

From the figures in Table 2, it appears that it is too strong to say that the introverted verb *wash* is “normally reflexive.” At most we can say that it is “commonly reflexive,” more commonly than *kill*. But what counts for explaining the coding of reflexive situations is the contrast between nonreflexive phoric pronouns and reflexive pronouns. The reason is that the nonreflexive phoric pronouns are the direct “competitors” of the reflexive markers, while non-phoric, full noun phrase objects are evidently non-coreferential and are, so to speak, outside the competition.

Thus, when a verb has a phoric notional object, in introverted verbs this is more commonly reflexive than nonreflexive. This is not so easy to see in English, because introverted verbs typically lack a phoric pronoun altogether when the notional object is coreferential with the subject. Some further data from Czech and German are shown in Table 3.

two *introverted* verbs:

			disjoint pronoun		reflexive pronoun	
German	waschen	'wash'	66	<b>(32%)</b>	141	<b>(68%)</b>
Czech	mýt, umýt, umývat	'wash'	28	<b>(22%)</b>	98	<b>(78%)</b>
German	verteidigen	'defend'	43	<b>(21%)</b>	162	<b>(79%)</b>
Czech	bránit	'defend'	7	<b>(4%)</b>	194	<b>(96%)</b>

two *extroverted* verbs:

			disjoint pronoun		reflexive pronoun	reciprocal pronoun	
German	hören	'hear'	196	<b>(96%)</b>	8	<b>(4%)</b>	0
Czech	slyšet	'hear'	201	<b>(98%)</b>	2	<b>(1%)</b>	2 1%
German	hassen	'hate'	160	<b>(76%)</b>	14	<b>(7%)</b>	37 18%
Czech	nenávidět	'hate'	104	<b>(76%)</b>	19	<b>(14%)</b>	13 10%

Table 3: Transitive verbs with coreferential and disjoint object pronouns (sources: for German: Cosmas Corpus of Institut für deutsche Sprache; for Czech: Czech National Corpus<sup>8</sup>)

So far I have presented the difference between introverted and extroverted verbs as a simple bifurcation. However, the quantitative perspective makes it clear that we are really dealing with a **continuous scale of increasing reflexive use**, with on the one hand verbs that are (almost) never used reflexively and on the other hand verbs that very frequently occur reflexively, with many different types in between. Strictly speaking, the prediction should therefore be:

<sup>8</sup> I am grateful to Sven Siegmund for help with the Czech and IDS Cosmas corpus counts.

(10) **Universal 1a**

In all languages, verbs with higher frequency of reflexive use show shorter reflexive-marking forms than verbs with lower frequency of reflexive use.

This makes further predictions about possible systems with more than two different reflexive-marking strategies (though these seem to be quite rare), and it is more easily testable than 9) when it is not fully clear which verbs belong to the introverted and extroverted classes. In Dutch, many verbs occur with both *zichzelf* and *zich*, and Smits et al. (2007) and Hendriks et al. (2008) have shown, using large corpora (up to 300 million words), that the use of *zich* strongly correlates with a high proportion of reflexive phoric pronouns, whereas the use of *zichzelf* correlates with a low proportion of reflexive phoric pronouns.<sup>9</sup>

An introversion/extroversion contrast has not just been observed for verbal actions, but also for adjectives (e.g. Zribi-Hertz 1995) and other types of predicates (e.g. Smith 2004). The following contrast from Zribi-Hertz (1995:347) is well known:

## (11) French

- a. Pierre<sub>1</sub> est fier de lui<sub>1/2</sub>.  
'Pierre is proud of himself/him.'
- b. Pierre<sub>1</sub> est jaloux de lui-même<sub>1/\*2</sub>. (...jaloux de lui<sub>2/\*1</sub>)  
'Pierre is jealous of himself.' ('...jealous of him...')

Zribi-Hertz limits herself to saying that there are two predicate types which must be lexically specified for [ $\pm$ disjoint reference]. This is probably correct as far as the grammar of French goes, but we can go further, because the contrast in (11) must reflect a universal tendency. There is a clear frequency asymmetry in corpus data, as shown in Table 4, using data from English. While a significant proportion of animate complement pronouns with *proud of* are reflexive (e.g. *proud of himself*), the reflexive occurrence of *jealous* is extremely rare and is not attested in the 100 million word British National Corpus.

	PERSONAL PRONOUN	REFLEXIVE PRONOUN		
<i>proud of</i>	212	<b>(84%)</b>	39	<b>(16%)</b>
<i>jealous of</i>	41	<b>(100%)</b>	0	<b>(0%)</b>

Table 4: Two adjectives with (animate) disjoint/coreferential pronoun complements (source: British National Corpus)

A lot more could be said about the diachronic mechanism by which discourse frequency/rarity is translated into shortness/length of coding, in addition to the basic insight that predictability allows shortness of coding, while nonpredictability requires explicitness of coding. However, I will not discuss the precise diachronic pathways in this article, which focuses on the correlation of reflexive-marking universals with usage regularities. The diachronic rise of overtly marked reflexives has been discussed elsewhere (especially for English; see Faltz 1985: ch. 4, König & Siemund 2000b, Levinson 2000:§4.4, Keenan 2003, Ariel 2008). More general questions about the diachronic rise of grammatical asymmetries in response to frequency asymmetries are addressed in Haspelmath (2008). I should perhaps emphasize that the functional explanation of the universals is essentially diachronic in nature, i.e. no claim is made that the usage frequencies are in any way relevant to the synchronic

<sup>9</sup> See also Miličević (2007) for an approach to reflexives (and reciprocals) that takes the continuum view seriously.



grammatical systems of languages. This approach in no way challenges the grammar-usage distinction (cf. Newmeyer 2003). What it challenges is the idea that language universals should in general be a direct consequence of the innate cognitive structures that make language acquisition possible.

#### 4. Length of the reflexive marker

The next universal to be discussed here concerns the complexity or length of reflexive markers as compared with the length of non-reflexive, disjoint-reference-marking phoric pronouns (see, e.g., Comrie 1999: 342, Levinson 2000:329, Mattausch 2007:§2.3).

##### (12) Universal 2

In all languages, the primary reflexive-marking strategy is at least as long as the primary disjoint-reference-marking strategy.

The notion *primary reflexive-marking strategy* is taken from Faltz (1985:4): It is the strategy used for subject-coreferential direct objects of extroverted transitive verbs, e.g. *herself* in *She admires herself*. The *primary disjoint-reference-marking strategy* is the strategy corresponding to English *him* in *They hate him*, i.e. the non-reflexive personal pronoun. Universal 2 thus says that expressions corresponding to *him* are never longer than expressions corresponding to *herself*. The universal is formulated in terms of “strategies” rather than (reflexive or disjoint) phoric expressions because the primary reflexive-marking strategy may be a verbal strategy, not using a referential expression.

Table 5 shows the reflexive markers and disjoint-reference markers of some languages where there is a difference between the two strategies.<sup>10</sup>

	REFLEXIVE-MARKING	DISJOINT-REFERENCE-MARKING
English	herself	her
Greek	ton eaftó tu	ton
Hebrew	(et) šacmo	oto
Turkish	kendini	onu
Oriya	nijaku	taaku
Lezgian	wič	am
Japanese	zibun (o)	Ø
Mandarin Chinese	zǐji	tā
German	sich	ihn
French	se	le
Swahili	ji-	mu-

Table 5: Reflexive markers and phoric disjoint-reference markers

The first seven languages provide evidence for the asymmetry noted in Universal 2 because the reflexive marker is longer than the nonreflexive phoric pronoun, whereas the last three languages do not provide evidence for it, though they are consistent with it. Languages where the nonreflexive phoric pronoun is longer than the primary reflexive marker are unattested.

This universal has not been widely discussed by theoretical linguists. But at least since Reinhart & Reuland (1993), the distinction between “complex anaphors” like English *herself* on the one hand, and “simplex-expression anaphors” like German *sich* or French *se* on the

<sup>10</sup> Languages may not make a distinction between reflexive-marking and disjoint-reference-marking elements (e.g. Loni, as illustrated in 6) above; see also Levinson 2000:338-341). Universal 2 makes no prediction about such languages.

other, has been prominent in the generative literature. This notion is not defined precisely by Reinhart & Reuland (1993:658), but if we assume that “complex” in “complex anaphor” means “morphologically complex,” then we can relate Reinhart & Reuland’s claims to Universal 2. One of their main principles (“Condition B”), attributed to the innate universal grammar, says: “A reflexive predicate is reflexive-marked,” where “reflexive-marked” means “lexically specified as reflexive” or “marked by means of a complex anaphor as one of its arguments” (Reinhart & Reuland 1993:663). Thus, Dutch *haat zichzelf* ‘hates herself’ is reflexive-marked by the complex anaphor *zichzelf*, and Dutch *wast zich* is reflexive-marked by lexical stipulation (the lexeme *wassen* ‘wash’ contains a lexical feature “reflexive”). Reinhart & Reuland therefore predict that with verbs that are not lexically reflexive, the reflexive marker should be “complex.”

However, this principle falls far short of subsuming Universal 2, being too restrictive and too permissive at the same time. On the one hand, it excludes cases like Oriya, Lezgian, Japanese and Mandarin, where the reflexive marker is synchronically simple from a morphological point of view. It is merely longer than the nonreflexive phoric,<sup>11</sup> thus providing evidence for Universal 2, and conflicting with Reinhart & Reuland’s Condition B (they admit this problem in their note 16, p. 667). The German and French cases also conflict with their Condition B, but do not contradict Universal 2.<sup>12</sup> On the other hand, Reinhart & Reuland allow the possibility of a language in which all verbs have two lexical entries, one of which is lexically reflexive, so that all verbs behave like English *shave*. Such a language does not seem to be attested, and Universal 2 correctly excludes it.<sup>13</sup>

The functionalist explanation of Universal 2 is that disjoint reference is overwhelmingly more frequent than coreference, so special coding is especially useful for coreference, and shorter coding for disjoint reference is economical. The data in Table 6 are from Ariel (2008: ch. 6).

disjoint	101	<b>(98%)</b>
coreferential	2	<b>(2%)</b>

Table 6. Coreferential and disjoint use of phoric object pronouns in transitive clauses (source: Ariel 2008:218, based on Santa Barbara Corpus of Spoken American English)

Thus, out of 103 pronominal referents 101 show disjoint reference, i.e. reflexive pronouns make up only 2% of all object pronouns. As we saw in the preceding section, reflexive use is rarer for extroverted verbs. Since the great majority of verbs are extroverted, it is clear that when all transitive verbs are considered, disjoint reference is overwhelmingly more frequent than coreference.

Faltz (1985:241-2) expressed this in the following terms: “In the case of a predication involving more than one argument, the unmarked situation is for the different arguments to have distinct referents.” Although Faltz does not say what exactly he means by “unmarked,” it seems fair to interpret it as a synonym of “most frequent” in this context (this sense of “unmarked” is common in the literature, cf. Haspelmath 2006). I am not aware of an explicit statement of the frequency-based economy explanation for Universal 2 in the earlier

<sup>11</sup> In Lezgian and Oriya, the reflexive is longer if one counts segments; if one counts syllables or moras, it is not necessarily longer, but it is not shorter either.

<sup>12</sup> In their note 16 (p. 667-668), Reinhart & Reuland suggest a modification of their theory, which could allow it to be extended to German, but which still fails to account for French.

<sup>13</sup> Mosel (1991) has observed that Samoan is a language that very rarely employs reflexive constructions using phoric pronouns. What is expressed by reflexive pronouns in English is expressed in a variety of very different ways in Samoan. Thus, there may well be languages lacking a primary reflexive-marking strategy (cf. Faltz 1985:18), but such languages do not contradict Universal 2.

functionalist literature, but it is very much in the spirit of Haiman's (1983) economic motivation.

However, Levinson (2000:328-9) has recently questioned the frequency-based explanation. He notes that "agents normally act upon entities other than themselves; the prototypical action—what is described by the prototypical transitive clause—is one agent acting upon some entity distinct from itself." This is fully in line with the frequentist account, as long as one interprets "normally" as "most frequently." But he continues:

If that is how the world stereotypically is, then an interpretation of an arbitrary transitive sentence as having referentially distinct arguments is given to us by the I-principle, which encourages and warrants an interpretation to the stereotype. Note that *this is not some kind of behaviorist presumption that the statistical preponderance of nonreflexive states of affairs, or even linguistic statements, is inductively learned and then reflected unwittingly in pragmatic presumption.*" [my emphasis]

In this passage, it is quite unclear what Levinson means by "how the world stereotypically is." Levinson does not define "stereotype," and he does not explain why it should be relevant for language how the world is. I maintain that the only thing that counts for a frequency-based explanation is frequency of use in human language (see §3). Levinson seems to regard statistical preponderances as something that cannot be relevant for theoretical considerations (cf. the use of the adjective "behaviorist"), but this is exactly what the Zipfian frequentist explanation claims: Inductive learning of statistical skewings in linguistic statements are reflected in the speakers' tendency to use explicit coding for the rarer situation, and this tendency then gets grammaticalized.

Mattausch (2007) has recently provided an account of Universal 2 in terms of bidirectional optimization, stochastic Optimality Theory, and iterated, bidirectional learning. His approach is fully compatible with my proposal here and provides further details on what must have been the diachronic trajectories that result in asymmetric coding.

## 5. Reflexive adnominal possessors

The third universal concerns the differentiation between reflexive and nonreflexive adnominal possessors. I am not aware of previous formulations of this universal in the literature (but Comrie (1999:338) hints at it).

### (13) Universal 3:

If a language uses a special reflexive pronoun for an adnominal possessor that is coreferential with the subject, then it also uses a special reflexive pronoun for the object, but not vice versa.

Thus, only three out of four logically possible language types are attested, as is illustrated in the table in (14), where for each attested type an exemplifying language is given.

(14)

		subject-coreferential pronouns in adnominal possessive position	
		normal	special reflexive
subject-coreferential pronouns in object position	special reflexive	English	Lezgian
	normal	Loni	—

In English, a special reflexive pronoun is used in object position, but the normal phoric pronoun is used in adnominal possessive position (as we saw earlier in (3a-b)). “Normal” here means the pronoun that is also used when the adnominal possessor is not coreferential with the subject.<sup>14</sup>

## (15) English

- a. She<sub>1</sub> killed herself<sub>1</sub>. (She<sub>1</sub> killed her<sub>2</sub>.)  
 b. She<sub>1</sub> killed her<sub>1/2</sub> lover. (\*She killed herself’s lover.)

In Lezgian, a special reflexive pronoun (*wič*) is used in case of subject-coreference, different from the normal phoric pronoun *am/ada-*:

## (16) Lezgian (see Haspelmath 1993:408-414)

- a. Alfija-di (wič-i) wič q’ena.  
 Alfija-ERG self-ERG self killed  
 ‘Alfija killed herself.’
- b. Alfija-di wič-in kic’ q’ena.  
 Alfija-ERG self-GEN dog killed  
 ‘Alfija<sub>1</sub> killed her<sub>1</sub> dog.’
- c. Alfija-di ada-n kic’ q’ena.  
 Alfija-ERG she-GEN dog killed  
 ‘Alfija<sub>1</sub> killed her<sub>2</sub> dog.’

And in Loniu, the normal phoric pronoun is used both in object position (as we saw in (6) above) and in adnominal possessive position:

## (17) Loniu (Hamel 1994:49)

- Hetow neʔəhin hetow tɔ ti/i tɔp a hetow.  
 3PCL girl 3PCL STAT weave basket POSS 3PCL  
 ‘The girls<sub>1</sub> are weaving their<sub>1/2</sub> baskets.’

Both the English type and the Lezgian type seem to be very widespread. The English type is also exemplified by Akan:

## (18) Akan (Faltz 1985:170-81)

- a. Mary hũũ nẽ hõ.  
 Mary see.PAST 3SG.POSS REFL  
 ‘Mary saw herself.’
- b. John praa nẽ ‘fie.  
 John sweep.PAST 3SG.POSS house  
 ‘John<sub>1</sub> swept his<sub>1/2</sub> house.’

<sup>14</sup> A reviewer points out that English has the expression *own*, which is sometimes taken to be a possessive reflexive (*She<sub>i</sub> killed her-own<sub>i</sub> lover*). However, *own* does not obey the same locality conditions as *(him-)self* (cf. *She thinks that her own lover did it*), and semantically it behaves more like the intensifier use of *(him-)self* (cf. König 2001:754-55).

The Lezgian type is also exemplified by Japanese:

(19) Japanese

- a. Ken wa zibun o seme-ta.  
 Ken TOP self ACC blame-PAST  
 ‘Ken blamed himself.’
- b. Jon<sub>1</sub> wa Marii<sub>2</sub> to zibun<sub>1/\*2</sub> no ie de hanasi o si-ta.  
 John TOP Mary with self GEN house in talk ACC do-PAST  
 ‘John had a talk with Mary in his/\*her house.’

Some languages allow both possibilities in adnominal possessive position, i.e. the reflexive pronoun (forcing a reflexive reading) or the nonreflexive pronoun (allowing a reflexive reading). Examples are Oriya and Tsez:

(20) Oriya (Ray 2000:588)

- a. Raama<sub>1</sub> (taa) nija-ku<sub>1</sub> bahut teke.  
 Rama his self-ACC much praises  
 ‘Rama praises himself very much.’
- b. Raama<sub>1</sub> nija<sub>1</sub> bahi paḍhilaa.  
 Rama self.GEN book reads  
 ‘Rama<sub>1</sub> reads his<sub>1</sub> book.’
- c. Raama<sub>1</sub> taa<sub>1/2</sub> bahi paḍhilaa.  
 Rama he.GEN book reads  
 ‘Rama<sub>1</sub> reads his<sub>1/2</sub> book.’

(21) Tsez (Polinsky & Comrie 1999: 329)

- a. ʃal-a nes-a nesi-r y<sup>ʃ</sup>utku r-oy-si.  
 Ali-ERG self-ERG self-DAT house GIV-make-PSTWIT  
 ‘Ali built a house for himself.’
- b. ʃal-a nes-a nesi-z qizaniyo-r y<sup>ʃ</sup>utku r-oy-si.  
 Ali-ERG self-ERG self-GEN2 family-DAT house GIV-make-PSTWIT  
 ‘Ali<sub>1</sub> built a house for his<sub>1</sub> family.’
- c. ʃal-a nesi-z qizaniyo-r y<sup>ʃ</sup>utku r-oy-si.  
 Ali-ERG he-GEN2 family-DAT house GIV-make-PSTWIT  
 ‘Ali<sub>1</sub> built a house for his<sub>1/2</sub> family.’

I am not aware of any contribution that the generative literature has made toward explaining this asymmetry. Chomsky’s (1981) binding theory is formulated in such a way as to predict the behavior in English, but it needs some amendments to allow for languages of the Lezgian or Japanese type. In any event, no generative explanation of the asymmetry noted in Universal 3 has become widely known.

But is a functionalist explanation possible? One obvious approach would be to claim that coreference of the adnominal possessor with the subject is not as surprising as coreference of the object with the subject, so that special marking of the coreferential possessor is less important than special marking of the coreferential object. But is coreference of the possessor

in any sense more “natural,” “(stereo)typical,” or “normal” than coreference of the object? Is (22a) “unmarked” or “stereotypical” compared to (22b)?

- (22) a. Robert<sub>1</sub> brought his<sub>1</sub> umbrella, so he<sub>1</sub> won't get wet.  
 b. Robert<sub>1</sub> has read his<sub>2</sub> book, so he<sub>1</sub> admires him<sub>2</sub>.

While Faltz and Levinson made their judgments about the special status of coreferential objects apparently on an intuitive basis, without statistical data, it seems more difficult to make an intuitive judgment in the case of (22a-b). But getting relevant statistical data is not difficult. Some are shown in Table 7.

### A. English *his*

(source: first 20 chapters of the English translation (CEV) of Genesis (the first book of the Bible))

subject-coreferential	43	(53%)	(Abraham went to his tent, Gen 18.6)
conjunct-coreferential	19	(23%)	(Noah and his sons, Gen 9.18)
disjoint	19	(23%)	(she was taken to his house, Gen 12.15)

### B. German *ihr*- ‘her; their’

(source: 19 of Grimm’s fairy tales)

subject-coreferential	79	(68%)
conjunct-coreferential	1	(1%)
disjoint	36	(31%)

Table 7. Coreferential and disjoint phoric possessors

Thus, it seems clear that adnominal possessive phoric pronouns are much more likely to be coreferential with the subject than object pronouns. As a result, they do not need special marking to the same extent as object pronouns, and thus they behave just like ordinary personal pronouns in many languages.

Now one might ask: But if possessive pronouns are more often than not coreferential with the subject, why do some languages have longer reflexive possessive pronouns than disjoint possessive pronouns (e.g. Japanese *zibun/kare*, Lezgian *wic#in/adan*)? This would seem to go against the spirit of Universal 2.

The answer is that in these languages, the possessive pronouns analogically follow the object pronouns. In Faltz’s (1985:118-119) terms, they exhibit **strategic streamlining** (i.e. possessive pronouns pattern after object pronouns), whereas English-type languages show **functional streamlining**. In other words: In some languages such as Japanese, “system pressure” beats economic motivation: The genitive *zibun no* is formed by analogy with the accusative *zibun o*. Note that strategic streamlining (= system pressure) can only create symmetries. It is still predicted that all asymmetries must be functionally motivated.<sup>15</sup>

If all cases where the longer reflexive pronouns in adnominal possessives are due to strategic streamlining or system pressure, we can still make a prediction about those cases where coreferential and disjoint possessive pronouns are differentiated, but the possessive reflexive is not “streamlined” (i.e. modeled on the non-possessive reflexive). In such cases, we expect the reflexive to be shorter, or at least not longer:

<sup>15</sup> However, we still need a fuller understanding of the directionality of system pressure. For example, we would not want to allow possessive pronouns to put pressure on object pronouns. More research is needed in this area.

**(23) Universal 3a**

If the possessive reflexive phoric ('his') does not form a coherent paradigm with the direct-object reflexive phoric ('himself'), it is at least as short as the non-reflexive possessive phoric.

I have no systematic evidence for the truth of this universal, but consider the following examples from Yimas and Finnish. Yimas, a Sepik language spoken in Papua New Guinea, has two different possessive phoric prefixes in the 3rd person singular, *m-* ("proximative") and *na-* ("obviative"). These are quite unrelated to the direct-object reflexive *panawt*, so they provide evidence for the asymmetry in Universal 3a. (In (24), G5 stands for gender number 5.)

**(24) Yimas (Foley 1991:180)**

- a.    *m-na-kn*                      *patn*                      *na-wayk-t*  
       3SG.PROX-POSS-G5.SG    *betelnut(G5)[SG]*    3SG.SUBJ-buy-PFV  
       'He<sub>1</sub> bought his<sub>1</sub> betelnut.'
- b.    *na-na-kn*                      *patn*                      *na-wayk-t*  
       3SG.OBV-POSS-G5.SG    *betelnut(G5)[SG]*    3SG.SUBJ-buy-PFV  
       'He<sub>1</sub> bought his<sub>2</sub> betelnut.'

Here the reflexive (or "proximative") possessive prefix *m-* shows a special short form, contrasting with the form *na-*, which also occurs as subject form on the verb.

In Finnish, phoric possessors are expressed by genitive personal pronouns and an agreeing possessive suffix on the possessed noun, e.g. *hän-en ruoka-nsa* [s/he-GEN food-3SG.POSS] 'his/her food'. When such a noun phrase occurs in object position, it is obligatorily interpreted as disjoint from the subject (cf. 25a). To obtain a subject-coreferential interpretation, a special short form is used that is obtained by omitting the personal pronoun (cf. 25b).

**(25) Finnish (Hennariikka Kairanneva, p.c.)**

- a.    *Hän*    *syö*    *hän-en*    *ruoka-nsa.*  
       she    eats    she-GEN    food-3SG.POSS  
       'She<sub>1</sub> eats her<sub>2</sub> food.'
- b.    *Hän*    *syö*    *ruoka-nsa.*  
       she    eats    food-3SG.POSS  
       'She<sub>1</sub> eats her<sub>1</sub> food.'

Again, the contrast in Finnish is quite unrelated to the direct-object reflexive *itse* 'himself', and it thus provides evidence for Universal 3a. (A similar situation is reported for Mapudungun by Smeets (2008:104).)

**6. Reflexives in locative phrases**

The fourth universal concerns the differentiation between reflexive and nonreflexive in locative phrases. This contrast was first discussed in some detail by Faltz (1985:§3.3), and the universal was formulated by Comrie (1999:338).

**(26) Universal 4:**

If a language uses a special reflexive pronoun in locative phrases, it also uses a special reflexive pronoun for objects, but not vice versa.

This is quite analogous to Universal 3: Special reflexive pronouns in locative phrases are less likely to exist than special reflexive pronouns in object position, just like reflexives as adnominal possessors. Again, only three out of four logically possible language types are attested, as is illustrated in the table in (27), where for each attested type an exemplifying language is given.

(27)

		subject-coreferential pronouns in locative position	
		normal	special reflexive
subject-coreferential pronouns in object position	special reflexive	English	German
	normal	Loniu	—

As in the possessive position, English shows no special reflexive pronoun in locative phrases:

(28) English: Maria<sub>1</sub> saw a snake near her<sub>1/2</sub>.

German is different in that it requires a special reflexive here (Faltz 1985:100).

(29) German: Maria<sub>1</sub> sah eine Schlange neben sich<sub>1</sub>/ihr<sub>2</sub>.

The third language type can again be exemplified by a language like Loniu, which completely lacks the contrast between reflexive and non-reflexive pronouns.

(30) Loniu (Hamel 1994:80)

Suʔu    ɲetu    suʔu    ime            pɛliŋɛʔi    suʔu.  
 3DU    child    3DU    3SG.come    with            3DU  
 ‘Their<sub>1</sub> two children<sub>2</sub> came to be with them<sub>1/2/3</sub>.’

In addition to Universal 4, which notes an asymmetry of distribution (like Universal 3), we also have some suggestive evidence for Universal 5, which notes an asymmetry of formal expression (like Universal 2).

**(31) Universal 5**

If different reflexive pronouns are used for objects and in locative phrases, the locative-phrase reflexive is phonologically less complex, other things being equal.

A version of this universal was noted by Faltz (1985:108).<sup>16</sup> A language that shows an interesting contrast relevant to Universal 5 is Dutch. Like English, Dutch differentiates between locative and object pronouns that corefer with the subject, but unlike English, it does not use the ordinary non-reflexive pronoun in locative constructions. Instead, it uses the “simplex-expression anaphor” *zich*, which is also used with introverted verbs (cf. Table 1

<sup>16</sup> In some languages, locative adpositions require independent pronouns, while the direct-object position can be filled with a clitic pronoun (e.g. French *il se voit* ‘he sees himself’ vs. *il voit un serpent derrière lui* ‘he sees a snake behind him’). In such languages, other things are not equal, and Universal 5 does not apply in the simple way, because independent pronouns are often longer than clitic pronouns.



above), whereas phoric pronouns in object position with ordinary (extroverted) verbs must be “complex anaphors” (*zichzelf*).

(32) Dutch (Reinhart & Reuland 1993:665-6)

- a. Max legt het boek achter zich.  
‘Max puts the book behind him.’
- b. Max haat zichzelf.  
‘Max hates himself.’

A generative explanation of this contrast has been provided by Reinhart & Reuland (1993), whose Condition B we already saw above (§4): “A reflexive predicate is reflexive-marked,” and “reflexive-marking” in their sense can only be achieved by a “complex anaphor.” Since, according to Reinhart & Reuland, locative phrases form their own predicates, they do not need to be (and in fact cannot be) “reflexive-marked,” so they do not occur with a complex reflexive pronoun.

However, different locative prepositions and different predicates behave differently, suggesting that this is not a matter of pure configurational syntax (predicate vs. no predicate). Faltz (1985:107) observes the following contrast:

- (33) a. Krag the robot placed a sandwich in front of him/?\*himself.  
b. Krag the robot unscrewed a panel in his abdomen and placed a sandwich inside himself/?him.

The occurrence of *himself* in (33b) (and marginally also in 33a) is explained by Reuland & Reinhart (1993) as due to the “logophoric” use of *himself*, which obeys mostly nonsyntactic conditions. What they cannot explain is that the simple *him* is questionable in (33b), and that, as Smith (2004: 598) notes, *himself* is sometimes quite impossible:<sup>17</sup>

- (34) He looked about him/\*himself.
- (35) She has a lot of money on her/\*herself.
- (36) The box has a spider in it/\*itself.
- (37) Dave put his past behind him/\*himself.

The functionalist explanation advocated here again observes that subject-coreference is significantly more common in locative phrases than with objects. Hence, languages do not need special reflexive pronouns as much as for objects, and reflexive pronouns can be shorter. Some frequency figures are given in Table 8.

---

<sup>17</sup> In English, the judgments for (28) and (33-37) vary quite a bit, and there are also some cases where the reflexive seems to be required in locative contexts (e.g. *John pulled the duvet over himself*, provided by a reviewer). Discussing the details of English usage is beyond the scope of this paper but clearly warrants further study.

## A. German locative prepositions

(source: Goethe Corpus of Institut für deutsche Sprache Mannheim)

bei sich	coreferential	93	(31%)
bei ihm/ihr/ihnen	disjoint	209	(69%)
vor sich	coreferential	188	(55%)
vor ihm/ihr/ihnen	disjoint	153	(45%)
hinter sich	coreferential	39	(48%)
hinter ihm/ihr/ihnen	disjoint	42	(52%)
unter sich	coreferential	30	(42%)
unter ihm/ihr/ihnen	disjoint	42	(58%)
über sich	coreferential	66	(47%)
über ihm/ihr/ihnen	disjoint	75	(53%)

## B. English locative prepositions

(source: British National Corpus, random selection of 50 occurrences):

near him	coreferential	10	(20%)
	disjoint	40	(80%)
behind him	coreferential	12	(24%)
	disjoint	38	(76%)
in front of him	coreferential	17	(34%)
	disjoint	33	(66%)
above him	coreferential	7	(14%)
	disjoint	43	(86%)
below him	coreferential	8	(16%)
	disjoint	42	(84%)

Table 8. Coreferential and disjoint use of phoric pronouns in locative phrases

These figures show that coreferential use of the pronoun is much more common in locative phrases with these prepositions than in object position (recall that Ariel found only 2% coreferential use of object pronouns). It is unclear why the coreferential use seems to be much less frequent in English than in German, but the figures seem robust enough to lend sufficient plausibility to the frequency-based explanation of Universals 4 and 5. Since coreferentiality between phoric pronoun and subject is much more likely in locative phrases than in direct-object position, there is less of a tendency to use special, longer reflexive pronouns in locatives.

## 7. Long-distance reflexives

The sixth and seventh universals concern long-distance reflexives. Again the relevant issues were first discussed by Faltz (1985:§3.6). The universals 6-7 are completely parallel to the universals 4-5 of the preceding section.<sup>18</sup>

### (38) Universals 6-7:

6. If a language uses a special reflexive pronoun in long-distance contexts, it also uses a special reflexive pronoun in local contexts, but not vice versa.
7. If a language has different reflexive pronouns in local contexts and long-distance contexts, the local reflexive pronoun is at least as complex phonologically as the long-distance reflexive.

Something like Universal 7 has long been widely discussed. Faltz (1985:153) observed that “compound reflexives tend to obey the [clause mate condition]” (i.e. to be limited to the same clause as the antecedent), and Pica (1987) claimed that long-distance reflexives are generally monomorphemic.

Some examples of languages that are consistent with these universals are given in Table 9.

	LOCAL REFLEXIVE	LONG-DISTANCE REFLEXIVE
Mandarin Chinese	(tā) ziji	ziji
Icelandic	sjálfan sig	sig
Dutch	zichzelf	zich
Telugu	tanu tanu	tanu
Bagvalal	e-b-da	e-b (Ljutikova 2001)
Malay	diri-nya	diri-nya
English	him-self	him-self

Table 9: Local reflexives and long-distance reflexives

A generative explanation for some of the effects of Universal 7 has been provided by Pica (1987), Cole et al. (1990), and others (see Cole et al. 2005 for a review). These authors argue that long-distance-reflexives become local by “head movement,” so that it is predicted that phrasal reflexives cannot occur in nonlocal contexts. Phrasal reflexives, which are necessarily polymorphemic, can only occur locally, whereas monomorphemic reflexives can be long-distance reflexives.

A serious conceptual problem with this proposal is that it has to claim that local reflexives like Dutch *zichzelf* are phrasal, in contrast to *zich*, which is not a phrase. The evidence for this view is minimal. An empirical problem is that counterexamples to Pica’s generalization that long-distance reflexives must be monomorphemic have been found: in Malay and English, the bimorphemic local reflexives are also used as long-distance reflexives (see also Huang 2000:96-7).<sup>19</sup> Since Universal 7 only claims that the local reflexives should be at least as complex, these languages do not constitute counterevidence to this universal.<sup>20</sup>

<sup>18</sup> Comrie (1999:338) formulates a universal that can be said to generalize over my Universals 3, 4, and 6: “all languages requiring reflexive pronouns at least somewhere have obligatory reflexive pronouns in this most local domain” (i.e. “that domain which includes only the arguments (subject and objects) of a single predicate”).

<sup>19</sup> However, Cole et al. (2005) claim that Malay and English are not counterexamples, because Faltz’s/Pica’s generalization applies only to long-distance reflexives that are bound anaphors, not to those that behave like pronominals (and Malay *dirinya* and English *himself* belong to this latter category).

<sup>20</sup> The Bagvalal long-distance reflexive *e-b*, which is bimorphemic as well, also contradicts the strictest interpretation of Pica’s generalization. However, its non-stem morpheme is an inflectional (gender) affix, so it obviously does not count as phrasal.

The functionalist explanation advocated here again appeals to frequency and economy. Phoric pronouns in subordinate clauses are much more likely to be (subject-) coreferential than phoric pronouns in clausemate object position. Hence they do not need as much coding as object pronouns, i.e. they can be identical to ordinary phoric pronouns, or they can be shorter than local reflexive pronouns.

The connection between the expectation of coreference and lack of specific overt coding was formulated very clearly by Comrie (1999:341): “As we move to more and more extended domains, the expectation of non-coreference is relaxed, so that ... at some particular point an individual language will decide to shift from reflexive to ordinary pronoun even in cases of coreference.” By “more and more extended domains” Comrie means the scale from the most local domain (comprising the predicate and its arguments), via the predicate’s adjuncts, to non-finite and finally finite subordinate clauses.

However, Comrie does not say why the expectation of non-coreference should be different in “more extended domains.” Once we look at this from the perspective of frequency of use, the answer becomes very simple: In long-distance contexts, non-coreference is less expected than in local contexts because it occurs far less often in actual discourse. In Table 10 I give some frequency figures from a small written corpus of German and a small spoken corpus of Czech. The data are limited to complement clauses corresponding to English *that* clauses.

#### A. German *dass*-clauses, all phoric pronouns

(source: German translations of Acts (Bible))

disjoint in the sentence	57	(47%)
coreferential with superordinate subject	46	(38%)
coreferential with superordinate nonsubject	14	(11%)
antecedent within subordinate clause	5	(4%)

#### B. Czech *že*-clauses, all phoric pronouns

(source: Czech National Corpus, sub-corpus of spoken language)

disjoint in the sentence	135	(55%)
coreferential with superordinate subject	76	(31%)
coreferential with superordinate nonsubject	15	(6%)
antecedent within subordinate clause	21	(9%)

Table 10. Coreferential and disjoint use of phoric pronouns in finite complement clauses

Thus, the statistical tendencies of phorics in complement clauses are very similar to phorics in adnominal possessive function and in locative phrases. It is therefore completely expected that their formal behaviour should show striking similarities.

### 8. Conclusion: Explanata and explanabilia

The universals corresponding to the first three contrasts in §1 have now been explained (plus a few more):

- (39) a. Russian \*Vanja nenavidit-sja. (Universal 1)  
 b. English \*Bob<sub>1</sub> saw him<sub>1</sub>. (Universal 2)  
 c. English \*Bob admires himself’s boss. (Universal 3)

But linguists are often (in practice, most of the time) also interested in language-particular facts. Can we also explain why Russian is not like German, or why English is not like Lezgian, for example? The answer is no (until we find further, hitherto unknown universals and explanations for them). Functional explanations are by their nature incapable of explaining language-particular facts, because the functional explanatory factors of frequency and economy have universal scope. We can say that the language-particular facts have been explained in a weak sense to the extent that they instantiate the universals (cf. Vennemann 1983), because knowing that a language-particular contrast falls under an explainable universal gives us an *Aha-Erlebnis* (relief from puzzlement). Explanation in the strong sense is possible only for grammatical universals (i.e. necessary properties of language), not for language-particular facts (i.e. accidental properties of language). That Russian is not like German is a historical accident, and we can no more explain the syntactic differences between individual languages than we can explain lexical differences like Russian *derevo* vs. German *Baum* (both mean ‘tree’).<sup>21</sup>

The fourth contrast of §1 has not been explained, and no attempt at explanation has been made, because there is no known universal that it instantiates. We simply do not know the cross-linguistic facts here.<sup>22</sup>

Thus, the functionalist frequency-based approach adopted here allows us to (strongly) explain a significant number of universals of reflexive marking, and to (weakly) explain facts of many languages that might initially be surprising. This approach makes many claims and predictions that are easily testable. Especially the claims about frequency distributions are easily testable by examining more corpora, and the predictions of the universals are easily testable because the universals make only minimal reference to controversial concepts.

A question that I have not addressed in this paper is what might be the source of the frequency asymmetries that we saw. This is an interesting question, but I assume that it is not relevant to explaining the grammatical universals at issue here. In principle one could imagine that the causal direction is the opposite from the direction proposed here, i.e. that the grammatical asymmetries are somehow the cause of the frequency asymmetries. Or one could imagine a factor that is simultaneously responsible for the frequency asymmetries and the grammatical asymmetries, so that frequency and grammar are not independent of each other, but both depend on a third factor. Both of these are logical possibilities, but neither has been advocated in the literature, probably for good reasons. The study of analogous phenomena in other areas of grammar (often under the rubric “markedness”) has shown that the set of factors that lead to frequency asymmetries in grammar is extremely diverse and disparate, but whenever there is a consistent frequency asymmetry, we get consistent results (see Haspelmath 2006:§4.2). Thus, while an answer to the question of why we find these frequency asymmetries will be instructive and provide an even deeper understanding of the underlying causes of the grammatical asymmetries, it is not a necessary ingredient to the present account, and the present account depends in no way on what the answer may turn out to be.

I conclude that the approach outlined in this paper has the virtue of avoiding bold speculative claims about speakers’ mental grammars, and bold speculative claims about the

<sup>21</sup> Linguists often try to push language-particular explanation further by constructing hyper-general language-particular descriptions, from which individual rules are said to follow. For example, one could try to come up with a hyper-general description of the contrast between *himself* and *him* in English that subsumes the contrasts in (3), (5) and (7). Such accounts, like Chomsky’s Binding Theory, are often ingenious and insightful (and they could certainly be called “explanations”) but they are very hard to falsify, and we have no way of knowing whether the speakers, and not just the linguists, make these generalizations.

<sup>22</sup> Ideally, the frequentist explanatory framework should allow us to predict universals from observed frequency asymmetries. This looks very difficult for this particular problem, because very large corpora would be required to find enough cases of partial coreference.

innate cognitive code (“Universal Grammar”).<sup>23</sup> Such claims are routinely made by generative approaches. This does not mean that my approach does not make bold claims. It does make some bold (and perhaps speculative) claims about universals of grammatical marking and universals of text frequency distributions. But these are easily falsifiable, unlike most generative proposals, which are notoriously hard to falsify.

## References

- Ariel, Mira. 2008. *Pragmatics and grammar*. Cambridge: Cambridge University Press.
- Chomsky, Noam A. 1981. *Lectures on government and binding*. Dordrecht: Foris.
- Calude, Andreea. 2007. Light and heavy reflexive marking: The middle domain in Romanian. *Annual review of Cognitive Linguistics* 5:239-269.
- Cole, Peter, Gabriella Hermon, and Li-May Sung. 1990. Principles and parameters of long-distance reflexives. *Linguistic Inquiry* 21:1-22.
- Cole, Peter and Gabriella Hermon. 2005. The typology of Malay reflexives. *Lingua* 115.5:627-644.
- Cole, Peter, Gabriella Hermon, and C.T. James Huang. 2005. Long distance anaphors: An Asian perspective. In Martin Everaert and Henk van Riemsdijk (eds.) *The Blackwell Companion to Syntax*. Malden, MA: Blackwell.
- Comrie, Bernard. 1989. *Language universals and linguistic typology*. 2nd ed. Oxford: Blackwell.
- , 1999. Reference-tracking: Description and explanation. *Sprachtypologie und Universalienforschung* 52:335-46.
- Everaert, Martin. 1986. *The syntax of reflexivization*. Dordrecht: Foris Publications.
- Faltz, Leonard M. 1985. *Reflexivization: A study in universal syntax*. New York: Garland.
- Foley, William A. 1991. *The Yimas language of New Guinea*. Stanford: Stanford University Press.
- Geniušienė, Emma. 1987. *The typology of reflexives*. Berlin: Mouton de Gruyter.
- Haiman, John. 1983. Iconic and economic motivation. *Language* 59: 781-819.
- Hamel, Patricia. 1994. *A grammar and lexicon of Loni, Papua New Guinea*. (Pacific Linguistics) Canberra: Australian National University.
- Haspelmath, Martin. 2004. Does linguistic explanation presuppose linguistic description? *Studies in Language* 28.3:554-579 (special issue guest edited by Martina Penke and Anette Rosenbach)
- , 2006. Against markedness (and what to replace it with). *Journal of Linguistics* 42.1:25-70.
- , 2008. Creating economical morphosyntactic patterns in language change. In Jeff Good, (ed.) *Language universals and language change*, 185-214. Oxford: Oxford University Press.
- , To appear. Framework-free grammatical theory. In Bernd Heine and Heiko Narrog (eds.) *The Oxford handbook of grammatical analysis*. Oxford: Oxford University Press.
- Hendriks, Petra Jennifer Spenader and Erik-Jan Smits, . 2008. Frequency-based constraints on reflexive forms in Dutch. *Proceedings of the 5th International Workshop on Constraints*

---

<sup>23</sup> Note that I do not claim that the cognitive code (“Universal Grammar”) plays no role in shaping grammars. That our innate predispositions are relevant for grammarians is clear from two simple considerations:

(i) Grammaticalization often leads to perfectly discrete rules, where the discourse asymmetry is much less crystal-clear (and anyway the very fact of grammaticalization shows that we must have a cognitive code for language -- otherwise no generalizations can be encoded);

(ii) Rules are known to be sensitive to categories of grammar, not to quantities (“grammars don’t count”). But clearly, the cognitive code is very permissive and allows many more grammars than are actually attested, just like the genetic code allows many more organisms than have a chance of survival (see Haspelmath 2004).

- and Language Processing (CSLP2008). Hamburg, Germany, August 11-15, 33-47. (<http://www.ruc.dk/upload/application/pdf/62445b39/rr122.pdf>)
- Huang, Yan. 2000. Anaphora. (Oxford Studies in Typology and Linguistic Theory) Oxford: Oxford University Press.
- Keenan, Edward. 2003. An historical explanation of some binding theoretic facts in English. In John Moore and Maria Polinsky (eds.) *The nature of explanation in linguistic theory*, 153-189. Stanford: CSLI Publications.
- König, Ekkehard and Peter Siemund. 2000a. Intensifiers and reflexives: A typological perspective. In Zygmunt Frajzyngier and Traci S. Curl. (eds.) *Reflexives: Forms and functions* (Typological studies in language, 40), 41-74. Amsterdam: Benjamins.
- , 2000b. The development of complex reflexives and intensifiers in English. *Diachronica* 17.1:39-84.
- König, Ekkehard. 2001. Intensifiers and Reflexives. In Martin Haspelmath, Ekkehard König, Wulf Oesterreicher and Wolfgang Raible (eds.) *Language typology and language universals*. Vol. 1, 747-60. Berlin: Mouton de Gruyter.
- König, Ekkehard and Vezzosi Letizia. 2004. The role of predicate meaning in the development of reflexivity. In Björn Wiemer, Walter Bisang and Nikolaus Himmelmann (eds.) *What makes Grammaticalization? A look from its fringes and its components*, 213–244. Berlin: Mouton de Gruyter.
- Langacker, Ronald W. 2000. A dynamic usage-based model. In Michael Barlow and Suzanne Kemmer (eds.) *Usage-Based models of language*, 1-63. Stanford: CSLI.
- Levinson, Stephen C. 2000. *Presumptive meanings: The theory of generalized conversational implicature*. Cambridge/MA: MIT Press.
- Ljutikova, Ekaterina A. 2001. *Anaforičeskie sredstva*. In A.E. Kibrik (ed.) *Bagvalinskij jazyk*, 615-681. Moskva: Nasledie.
- Mattausch, Jason. 2007. Optimality, bidirectionality and the evolution of binding phenomena. *Research on Language and Computation* 5:103-131.
- Miličević, Maja. 2007. *The acquisition of reflexives and reciprocals in L2 Italian, Serbian and English*. Ph.D. dissertation, University of Cambridge.
- Miller, Amy W. 2001. *A Grammar of Jamul Tiipay*. Berlin: Mouton de Gruyter.
- Mosel, Ulrike. 1991. Transitivity and reflexivity in Samoan. *Australian Journal of Linguistics* 11:175-194.
- Newman, Paul. 2000. *The Hausa language: An encyclopedic reference grammar*. New Haven: Yale University Press.
- Newmeyer, Frederick J. 2003. Grammar is grammar and usage is usage. *Language* 79: 682-707.
- Pica, Pierre. 1987. On the nature of the reflexivization cycle. *North-Eastern Linguistics Society* 17:483-499. (University of Massachusetts, Amherst, GLSA)
- Polinsky, Maria and Bernard Comrie. 1999. Reflexivity in Tsez. In E.V. Raxilina and Ja. G. Testelec (eds.) *Tipologija i teorija jazyka: ot opisanija k ob'jasneniju*, 319-339. Moskva: Jazyki ruskoj kul'tury.
- Postal, Paul. 1970. The method of Universal Grammar. In Paul Garvin (ed.) *On method in linguistics*, 113-131. The Hague: Mouton.
- Ray, Tapas S. 2000. Lexical anaphors and pronouns in Kannada. In Barbara C. Lust *et al.* (eds) *Lexical anaphors and pronouns in selected South Asian languages*, 575-636. Berlin: Mouton de Gruyter.
- Reinhart, Tanya and Eric Reuland. 1993. Reflexivity. *Linguistic Inquiry* 24:657-720.
- Smeets, Ineke. 2008. *A grammar of Mapuche*. Berlin: Mouton de Gruyter.
- Smith, Mark. 2004. Light and heavy reflexives. *Linguistics* 42.3:573-615.

- Smits, Erik-Jan, Petra Hendriks and Jennifer Spenader. 2007. Using very large parsed corpora and judgement data to classify verb reflexivity. In A. Branco (ed.) *Anaphora: Analysis, algorithms and applications*. Selected papers from the 6th Discourse Anaphora and Anaphor Resolution Colloquium, DAARC 2007, Lagos, Portugal, March 2007. (Lecture Notes in Artificial Intelligence 44410), Springer, 77-93. ([http://www.ai.rug.nl/~spenader/public\\_docs/Daarc2007Smitsetal.pdf](http://www.ai.rug.nl/~spenader/public_docs/Daarc2007Smitsetal.pdf))
- Vennemann, Theo. 1983. Causality in language change: Theories of linguistic preferences as a basis for linguistic explanations. *Folia Linguistica Historica* 4:5-26.
- Zipf, George K. 1935. *The psycho-biology of language: An introduction to dynamic philology*. Houghton Mifflin (Republished 1965 by MIT Press.)
- Zribi-Hertz, Anne. 1995. Emphatic or reflexive? On the endophoric character of French lui-même and similar complex pronouns. *Journal of Linguistics* 31:333-74.

Author's contact information:

Martin Haspelmath  
Max Planck Institute for Evolutionary Anthropology  
Deutscher Platz 6  
D-04103 Leipzig, Germany  
[haspelmath@eva.mpg.de](mailto:haspelmath@eva.mpg.de)