

## Some universals of grammar with particular reference to coding asymmetries (2)

Martin Haspelmath

### 1. Universal asymmetric coding patterns

(cf. Haspelmath 2020a)

- (1) a. Greenberg's Universal 35  
There is no language in which the plural does not have some nonzero allomorphs, whereas there are languages in which the singular is expressed only by zero. The dual and the trial are almost never expressed only by zero.
- b. Greenberg's Universal 38  
Where there is a case system, the only case which ever has only zero allomorphs is the one which includes among its meanings that of the subject of the intransitive verb.

	Hebrew	Khanty
SG	<i>yom</i>	<i>xot</i>
PL	<i>yam-im</i>	<i>xot-ət</i>
DL	<i>yom-ayim</i> 'day(s)'	<i>xot-ɣən</i> 'house(s)'

	English	German	Quechua
NOM	<i>he</i>	<i>Herr Kim</i>	<i>wasi</i> 'house'
ACC	<i>hi-m</i>	<i>Herr-n Kim</i>	<i>wasi-ta</i>

cf. Greenberg's (1966): grammatical number, gender, person, case, tense, aspect, mood and other domains often show asymmetries which he described with the terms "marked/unmarked".

In general, a **universal coding asymmetry** involving construction type 1 (e.g. singular) and construction type 2 (e.g. plural) implies the universal claim in (2).

- (2) If a language has an asymmetric coding contrast between construction type 1 (universally more frequent) and construction type 2 (universally less frequent), then construction type 1 shows a strong tendency to be coded with a shorter shape than construction type 2, and often by zero.

But not all languages and all grammatical contrasts show asymmetric coding. Languages may have **symmetric coding**, where either both constructions are equally coded, or both are left uncoded. For the simple case of singular and plural, these two cases can be illustrated by (3a-b).

- (3) a. Modern Greek (symmetric overt)

SG	<i>viólí-o</i> 'book'
PL	<i>viólí-a</i> 'books'

b. Mandarin Chinese (symmetric zero)

SG	<i>shū</i>	'book'
PL	<i>shū</i>	'books'

In languages with symmetric coding, a competing constraint takes precedence:

- Modern Greek: the general preference to express grammatical meanings **explicitly**
- Mandarin Chinese: the general preference to save coding energy and to leave inferrable meanings unexpressed.

English-type languages:	EFFICIENT coding system at the price of asymmetric or non-uniform coding
Greek-type languages:	non-efficient but UNIFORMLY EXPLICIT coding
Mandarin-type languages:	non-efficient but UNIFORMLY PARSIMONIOUS coding

All three language systems are optimal in their own way (the form-frequency prediction is relevant only to cases where the coding is asymmetric).

The implicational universal in (2) predicts that **an asymmetric counter-efficient pattern** does not exist (where the singular has an overt marker but the plural is left uncoded, as in the Pseudo-Greek pattern in (4)).

(4) (Pseudo-Greek, hypothetical)

SG	<i>vivlí-o</i>	'book'
PL	<i>vivlí</i>	'books'

Why is the functional-adaptive explanation proposed at the level of language universals, not at the level of particular languages?

For example, English has two future-tense constructions, with *will* and with *going to* (or *gonna*).

These are slightly different semantically, and they are formally asymmetric (*will* is shorter than *gonna*, 3 vs. 4 segments). Is it claimed that **this is because the *will*-future is more frequent than the *gonna*-future?**

NO: I make no predictions about such cases, because functional-adaptive explanations only work at the population level (i.e. they explain tendencies found in populations of languages).

Language histories are subject to a large number of contingencies, and the adaptive forces are relatively weak. The *gonna*-future could become popular very quickly for social reasons and thus more frequent than the *will*-future (this may well have happened in some varieties of English). But it would not be an exception to a general trend, because there is no general trend for languages to have two different future tenses of this sort.

## 2. Differential coding

Form-frequency correspondences are also widely found in situations where a single grammatical meaning is coded **differently in different grammatical contexts or in different lexical subclasses**: DIFFERENTIAL CODING.

e.g. differential object marking (or differential P flagging):

- (5) a. *Veo la casa.*  
I.see the house.  
'I see the house.'
- b. *Veo a la mujer.*  
I.see ACC the woman.  
'I see the woman.'
- (6) a. *Kani-ti sefer.*  
bought-1SG book  
'I bought a book.'
- b. *Kani-ti et ha-sefer.*  
bought-1SG ACC the-book  
'I bought the book.'

– the patient meaning is expressed in two different ways, depending on whether the P nominal is inanimate or animate (in Spanish), or indefinite or definite (in Hebrew).

– the frequency claim for Spanish concerns the **relative frequency of the patient meaning within the sets of all inanimate nominals and all animate nominals**:

The claim that inanimate Ps are more frequent than animate Ps means that a greater proportion of inanimate nominals have the P role than of animate nouns. (The analogous situation holds for Hebrew indefinite and definite Ps.)

e.g. agent-patient coreference:

Sometimes, grooming (or INTROVERTED) verbs like 'wash' or 'shave' show different behaviour from other-directed (or EXTROVERTED) verbs like 'kill' or 'hate', as seen in English (in 7) and Russian (in 8) (Haiman 1983: 803; Geniušienė 1987).

- (7) a. *They shaved.* (= 'They shaved themselves.')
- b. *They hate themselves.*
- (8) a. *Oni myli-s'.*  
they washed-REFL  
'They washed.'
- b. *Oni ubili sebja.*  
they killed themselves  
'They killed themselves.'

In these examples, reflexive marking is differential in that it depends on the lexical subclass (introverted vs. extroverted).

The shorter forms (zero-coded in English, reflexive *-s'* in Russian) are used with introverted verbs, because these occur more frequently with agent-patient coreference than extroverted verbs.

More generally, in a differential-coding pair, we are dealing with **a USUAL ASSOCIATION of a grammatical meaning with a grammatical context or a lexical subclass**. The claim is that such usual associations need less coding than unusual associations.

### 3. Differential-coding pairs in the nominal domain

#### 3.1. Accusative marking on inanimate vs. animate P-arguments

In many languages, P-arguments are coded differentially when they are animate, and as far as I know, the animate P always has a marker, whereas the inanimate P is uncoded (Bossong 1985; 1991; Iemmolo 2014).

	Spanish	Armenian
INANIMATE	Ø <i>la casa</i> 'house'	<i>mek</i> 'another one (inanimate)'
ANIMATE	<i>a la mujer</i> 'woman'	<i>mek-i</i> 'another one (animate)'

#### 3.2. Ergative marking on 1st/2nd person pronouns vs. full nominals

In some languages, A-arguments are coded differentially when they are 1st or 2nd person (locuphoric), and this generally means that they lack an overt ergative marker, in contrast to other kinds of arguments, especially full nominals (Dixon 1994: 86).

	Dyirbal	Georgian
1ST PERSON PRONOUN	<i>ɲadya-Ø</i>	<i>me-Ø</i>
FULL NOMINAL	<i>yarra-ɲgu</i> 'man'	<i>mama-m</i> 'father'

#### 3.3. Locative marking on place names vs. inanimate nouns vs. animate nouns

In a substantial number of languages, locative flagging is differential, such that place names tend to have the shortest coding (often zero, as in Tswana), and animate nouns tend to have the longest coding, with inanimate nouns intermediate between the two (Aristar 1997; Creissels & Mounole 2011; Haspelmath 2019c; Michaelis 2019).

	Basque	Tswana	Tamil
PLACE NAME	<i>Bilbo-n</i> 'in Bilbao'	<i>Gaborone</i> 'at Gaborone'	
INANIMATE	<i>mendi-tan</i> 'at the mountain'	<i>toporo-ng</i> 'in town'	<i>N-il</i>
ANIMATE	<i>neska-rengan</i> 'at the girl's'		<i>N-itam</i>

Again, this is because place names are usually associated with locative use, while this is less common for inanimate nouns, let alone animate nouns. The parallels with differential object marking are so striking that Haspelmath (2019c) calls this pattern "differential place marking".

### 3.4. Plurative and singulative marking on individualist vs. gregarious nouns

In some languages, there is a subclass-conditioned difference in singular and plural coding, such that “individualist nouns” (those that tend to occur with uniplex meaning) have overt plural (plurative) marking, while “gregarious” nouns (those that are usually associated with multiplex meaning) have singulative (overt singular) marking (*-en* in Welsh, *-ina* in Russian). For detailed discussion, see Haspelmath & Karjus (2017) and Grimm (2018).

		Welsh		Russian		English
INDIVIDUALIST	MULTI	<i>cath-od</i>	‘cats’	<i>koty</i>	‘cats’	<i>cat-s</i>
	UNI	<i>cath-Ø</i>	‘cat’	<i>kot-Ø</i>	‘cat’	<i>cat-Ø</i>
GREGARIOUS	MULTI	<i>moron</i>	‘carrots’	<i>kartofel’</i>	‘potatoes’	<i>salt</i>
	UNI	<i>moron-en</i>	‘carrot’	<i>kartofel-ina</i>	‘potato’	<i>grain of salt</i>

### 3.5. Adpossession marking with inalienable vs. alienable nouns

In languages where inalienable (kinship and body-part term) nouns behave differently from alienable nouns in adpossession constructions, the possessive markers are shorter (and typically zero) for the inalienable subclass, because of the usual association of the adpossession meaning with kinship and body-part nouns (Haspelmath 2017).

	Maltese		Jeli	
INALIEN	<i>id-Ø-i</i>	‘my hand’	<i>Soma Ø buloni</i>	‘Soma’s arms’
ALIEN	<i>il-ktieb tiegh-i</i>	‘my book’	<i>Soma ra monbilo</i>	‘Soma’s car’

### 3.6. Definiteness marking with vs. without possessor

Some languages have differential definiteness marking, omitting the definite article in the presence of an adpossession nominal (Haspelmath 1999).

	German	Hebrew		Welsh	
POSSESSED	<i>mein Ø Buch</i>	<i>Ø-sifr-i</i>	‘my book’	<i>Ø car y meddyg</i>	‘the doctor’s car’
UNPOSSESSED	<i>das Buch</i>	<i>ha-sefer</i>	‘the book’	<i>y car</i>	‘the car’

## 4. Differential-coding pairs in the predicational domain

### 4.1. Reflexive marking on introverted vs. extroverted verbs

Some languages have subclass-conditioned differential marking of reflexivity, with introverted verbs (those that are usually associated with coreferential objects) showing shorter coding than extroverted verbs (those that tend to have noncoreferential anaphoric objects) (Haiman 1983: §1.2.2; Haspelmath 2008c).

	Russian	Dutch	Greek	English
GROOMING	<i>moet-sja</i>	<i>wast zich</i>	<i>plen-ete</i>	<i>he washes Ø</i>
EXTROVERTED	<i>vidit sebja</i>	<i>ziet zichzelf</i>	<i>vlép-i ton eaftó tu</i>	<i>he sees himself</i>

## 4.2. Addressee and aliophoric person marking in imperatives vs. indicatives

There is a usual association of person with mood, in that imperatives tend to have second person subjects and indicatives tend to have 3rd person subjects. Languages often have split coding of bound 2nd and 3rd person person forms, and as predicted, there is a tendency for 2nd person form to be short or zero in imperatives (Aikhenvald 2010: 46), and 3rd person forms to be short or zero (Siewierska 2010).

		Latin		Turkish	
IMPERATIVE	2ND	<i>lauda-Ø</i>	'(you) praise!'	<i>bak-Ø</i>	'(you) look'
	3RD	<i>lauda-to</i>	'let her praise!'	<i>bak-sın</i>	'let her look'
INDICATIVE	2ND	<i>lauda-v-isti</i>	'you praised'	<i>bak-ıyor-un</i>	'you are looking'
	3RD	<i>lauda-v-it</i>	'she praised'	<i>bak-ıyor-Ø</i>	'she is looking'

## 4.3. Attributive and predicative marking on property roots (adjectives) vs. action roots (verbs)

Property concept roots tend to occur in attributive discourse function, while action roots are usually associated with predicative function.

Many languages have a split in how they treat content roots:

- property concept roots have short (or Ø) coding in attributive function
- action roots tend to need overt attributive markers  
(participial affixes or relativizers).

By contrast, in predicative function, property concept roots tend to need special marking by a copula, while action roots do not (Croft 1991: 67).

		German	French	M. Chinese	
ATTR	PROPERTY	<i>klein-Ø-es Kind</i>	<i>petit Ø enfant</i>	<i>xiǎo Ø háizi</i>	'small child'
	ACTION	<i>spiel-<b>end</b>-es Kind</i>	<i>enfant <b>qui</b> joue</i>	<i>wán <b>de</b> háizi</i>	'child who plays'
PRED	PROPERTY	<i>das Kind <b>is-t</b> klein</i>	<i>l'enfant <b>est</b> petit</i>	<i>háizi Ø xiǎo</i>	'the child is small'
	ACTION	<i>das Kind spiel-Ø-t</i>	<i>l'enfant joue-Ø</i>	<i>háizi Ø wán</i>	'the child plays'

## 5. Scales of differential coding domains

### 5.1. Scale for causative marking

**spontaneity scale** (Haspelmath 2016: 34):

transitive ('cut') > unergative ('run') > automatic ('freeze (intr.)') > costly ('break')

**Table 1. Five types of verb meanings on the spontaneity scale: Some examples**

<b>transitive</b> <b>(most spontaneous)</b>	<b>unergative</b>	<b>automatic</b>	<b>unaccusative</b> <b>costly</b>	<b>agentful</b> <b>(least spontaneous)</b>
'cut', 'wash', 'throw', 'eat', 'hit', 'see'	'talk', 'dance', 'walk', 'play', 'work', 'scream'	'melt', 'freeze', 'dry', 'wake up', 'sink', 'go out (fire)'	'break (intr.)', 'split (intr.)', 'open (intr.)', 'close (intr.)', 'change (intr.)', 'gather (intr.)'	'be cut', 'be washed', 'be thrown', 'be eaten', 'be hit', 'be seen'

**Table 2. Synthetic causatives: Some language types**

(sC = short synthetic causative, lgC = long synthetic causative, anaC = analytic)

	<b>transitive</b> (‘cut’)	<b>unergative</b> (‘talk’)	<b>unaccusative</b> (‘freeze; break’)
Tuvan	C	C	C
Indonesian	(anaC)	C	C
O’odham	(anaC)	(anaC)	C
Georgian	lgC	sC	sC
Halkomelem	lgC	lgC	sC
(unattested)	sC	lgC	anaC
(unattested)	lgC	anaC	sC
(unattested)	anaC	sC	lgC

- (9) automatic verb meanings
- |              | noncausal<br>(basic verbs)                              | causal<br>(causative verbs)   |
|--------------|---|---|
| a. ‘melt’    | French<br>Arabic<br>Hindi-Urdu<br>Hungarian<br>Lezgian  | <i>fondre</i><br><i>saaha</i><br><i>pighal-</i><br><i>olvad-</i><br><i>c’ur-un</i>      |
| b. ‘wake up’ | Finnish<br>Lithuanian<br>Mongolian<br>Turkish<br>Udmurt | <i>herätä</i><br><i>pabus-ti</i><br><i>ser-ex</i><br><i>uyan-mak</i><br><i>sajka-ny</i> |
- (10) costly verb meanings
- |            | noncausal<br>(anticausative verbs)                    | causal<br>(basic verbs)   |
|------------|---|---|
| a. ‘break’ | Armenian<br>Hebrew<br>Japanese<br>Romanian<br>Turkish | <i>yard-el</i><br><i>ni-šbar</i><br><i>war-e-ru</i><br><i>se rupe</i><br><i>kır-ıl-mak</i>  |
| b. ‘open’  | Arabic<br>Finnish<br>French<br>Swahili<br>Udmurt      | <i>fataha</i><br><i>avautua</i><br><i>s’ouwir</i><br><i>fungu-k-a</i><br><i>ust-išky-ny</i> |

**Table 5. Causatives and anticausatives: Some language types**  
(**caus** = causative, analytic or synthetic, **anticaus** = anticausative)

	transitive (‘cut’)	unergative (‘talk’)	automatic (‘freeze’)	costly (‘break’)
Russian	caus	caus	anticaus	anticaus
Udmurt	caus	caus	caus	anticaus
Indonesian	caus	caus	caus	caus
(unattested)	caus	caus	anticaus	caus
(unattested)	caus	anticaus	caus	caus

## 5.2. Scale for adpossessive marking

**possessor-prominence scale** (Haspelmath 2017: 219):

super-alienable (‘sky’) > alienable (‘house’) > inalienable (‘stomach’)

Alienable nouns tend to require a special marker for possessed use (e.g. a genitive or a construct marker), because it is unexpected to hear about a possessed alienable thing concept. (Alienable things are very likely to occur unpossessed.)

		nonpossessed		possessed	
possessive	O’odham	<i>mi:stol</i>	‘cat’	<i>ñ-mi:stol-ga</i>	‘my cat’
<b>anti</b> possessive	Huastec	<i>čuk-lek</i>	‘stomach’	<i>ʔu-čukul</i>	‘my stomach’

By contrast, inalienable nouns (denoting kinship and body-part concepts) tend to require a special (**antipossessive**) marker for nonpossessed use, because it is less expected to hear about a nonpossessed use with an inalienable noun. (Such nouns are not likely to occur without a possessor.)

## 5.3. Scale for plurative marking

**individuation scale** (cf. Grimm 2018):

individuals (‘house’) > aggregates (‘peas’) > substances (‘water’)

Individual nouns tend to require a special plural marker for multiplex use, because it is unexpected to hear about a multiplex individualist thing. (Individualist thing concepts are very likely to occur as uniplexes.)

		uniplex		multiplex	
plural	English	<i>house</i>		<i>hous-es</i>	
<b>anti</b> plural	Welsh	<i>plent-yn</i>	‘child’	<i>plant</i>	‘children’

By contrast, aggregate nouns (small animals, fruits, children) tend to require a special (**singulative**) marker for uniplex use, because it is less expected to hear about a uniplex use of a gregarious noun. (Such nouns are not likely to occur as uniplexes.)

#### 5.4. Scales for argument flagging

(cf. Haspelmath 2020b)

- a. inherent prominence  
 person scale: locuphoric (1st/2nd) > allophoric (3rd person)  
 (full) nominality scale: person form (independent or index) > full nominal  
 animacy scale: human (> animal) > inanimate
- b. discourse prominence  
 specificity scale: definite (> specific indefinite) > indefinite nonspecific  
 givenness scale: discourse-given > discourse-new  
 focus scale: background > focus

#### USUAL ASSOCIATIONS:

##### (11) single-argument association tendencies

- a. the A and the R tends to be referentially prominent  
 b. the P and T tends to be referentially non-prominent

##### (12) scenario association tendencies

- a. the A tends to be referentially more prominent than the P  
 b. the R tends to be referentially more prominent than the T

##### (13) The single-argument coding universal

If a language has an asymmetric single-argument coding split, then the coding is longer for prominent P/T-arguments or for non-prominent A/R-arguments.

e.g. **differential object (P) flagging**: givenness-conditioned, person-conditioned:

##### (14) Persian (Dalrymple & Nikolaeva 2011: 108-112)

- a. *man ketâb-râ xarid-am.*  
 I book-ACC buy.PST-1SG  
 'I bought the book.'
- b. *man sib-i(\*-râ) xord-am.*  
 I apple-INDF(-ACC) eat.PST-1SG  
 'I saw an apple.' (accusative flag is not allowed on nontopical P)
- c. *ki mašin-i-\*(râ) did?*  
 who car-INDF-(ACC) see.PST[3SG]  
 'Who saw a car?' (accusative flag is required on topical P)

##### (15) Abruzzese (dialect of Arielli; D'Alessandro 2017)

- a. *So vistə a mme / a tte.*  
 be.1SG seen to me / to you  
 'I have seen myself / you.'
- b. *Semə vistə a nnu / a vvu.*  
 be.1PL seen to us / to you  
 'We have seen us / you.'
- c. *\*So vistə a Marije / a jissə / a quillə.*  
 be.1SG seen to Maria / to them / to them  
 ('We have seen Maria / them.')

**differential recipient (R) flagging:** nominality-conditioned, specificity-conditioned

(16) Northeastern Neo-Aramaic of Telkepe (Coghill 2010: 226-228)  
(full nominal R)

a. *wəl-lə pāra ta xa-məskenn*  
gave-he money to a.certain-poor.person  
'He gave money to a certain poor person.' (= Coghill's 11b)

b. *kəm-yāwəl-lə ta məskenn*  
PST-he.give-3SG.M.OBJ to poor.person  
'He gave it to a poor person.' (= 14b)

(person-form R)

c. *kəm-yāwəl-lə hadiyn*  
PST-he.give-3SG.M.OBJ present  
'He gave him a present.' (= 14c)

d. *kəm-yāwəl-lux-ila*  
PST-he.give-2SG.M.OBJ-3SG.M.SEC  
'He gave them to you.' (= 16b)

(17) Wolof (Atlantic; Becher 2005: 19)

a. *Jox naa xale bu jigéen ji benn velo.*  
give 1SG girl DEF INDF bicycle  
'I gave the girl a bicycle.'

b. \**Jox naa benn xale bu jigéen velo bi.*  
give 1SG INDF girl bicycle DEF  
'I gave a girl the bicycle.'

c. *Jox naa velo bi ci benn xale bu jigéen.*  
give 1SG bicycle DEF to INDF girl  
'I gave the bicycle to a girl.'

(18) **The scenario universal**

If a language has an asymmetric scenario split, then the coding is longest for **upstream scenarios**, shortest for **downstream scenarios**, and intermediate for balanced scenarios.

upstream scenario: a scenario where the higher-ranked argument is less prominent (e.g. 3>1, inanim>anim)

downstream scenario: a scenario where the higher-ranked argument is more prominent (e.g. 2>3, def>indef)

**Person-conditioned special P flagging.** In Kolyma Yukaghir, a special accusative flag is required on the P when the A is allophoric (cf. Keine 2010: §6.3).

- (19) Kolyma Yukaghir (Russian Far East; Maslova 2003: 89; 10)
- a. *met es'ie tet pulut-kele kudede-m*  
 my father.NOM your husband-ACC kill-TR.3SG  
 'My father has killed your husband.' (A is 3rd person, special accusative flag)
- b. *met tolow kudede*  
 I.NOM deer.NOM kill.TR.1SG  
 'I killed a deer.' (A is locuphoric, no P flagging)

**Person-conditioned special A flagging.** In Sahaptin, a special ergative flag is required on the A when the P is locuphoric (cf. Keine 2010: §6.1)

- (20) Sahaptin (Sahaptian; Pacific Northwest) (Rude 2009: 13-14)
- a. *ku =š i-q'ínun-a tílaaki-nim*  
 and1SG 3.NOM-see-PST woman-ERG  
 'And the woman saw me.' (P is locuphoric, special ergative flag)
- b. *ku i-q'ínun-a áswan-Ø tílaaki-na*  
 and3.NOM-see-PST boy-ABS woman-ACC  
 'And the boy saw the woman.' (P is 3rd person, no A flagging)

### Special T coding conditioned by nominality of R

While the use of a special R marker to code the person-upstream scenarios  $3 > 1$  and  $3 > 2$  is perhaps the most widespread pattern, some languages use special focal forms of T when the R is a person form.

For example, Modern Greek has a set of Genitive (i.e. dative) and Accusative proclitics used in downstream and allophoric balanced scenarios, as seen in (21a). But these proclitics cannot be used in upstream scenarios, as seen in (21b).

- (21) a. *Tu to éðose.*  
 him.DAT it.ACC he.gave  
 'He gave it to him.' ( $3 > 3$ , person-balanced)
- b. \**Tu me éðose.*  
 him.DAT me.ACC he.gave  
 ('He gave me to him.') ( $3 > 1$ , person-upstream)
- c. OK: *Tu éðose eména.*  
 him.DAT he.gave me.FULL.ACC  
 'He gave me to him.' (Lit. 'He gave ME to him.')

The following universal is even more general:

### (22) The role-reference association universal

Deviations from usual associations of role rank and referential prominence tend to be coded by longer grammatical forms.

It also covers passive constructions...

## Appendix: Some illustrative frequency figures

This appendix gives some illustrative figures from the earlier literature to show that it is very plausible that Universal 2 (§2) is true: “Arguments with higher-ranked roles (A, R) tend to be more referentially prominent than arguments with lower-ranked roles (P, T), and vice versa.” These figures are only from seven different languages, as frequency figures are not often given in the literature, and I did not do any corpus research myself. But they should suffice for initial plausibility.

### Animacy of A and P

Swedish (Dahl & Fraurud 1996: 51)

	human	nonhuman	total
A	1766	1361	3127
P	580	3896	4476

Movima (Bolivia; Haude 2014: 9)

	human	nonhuman	total
A	896	355	1251
P	370	885	1255

### Full nominality of A and P

Vera'a (Vanuatu; Haig & Schnell 2016: 599)

	person form	full nominal	total
A	680	115	795
P	325	580	905

### Specificity of A and P

English (Jäger 2007: 80, Table 3)

	definite	indefinite	total
A	973	32	1005
P	467	538	1005

### Person of A and P

English (Jäger 2007: 80, Table 3)

	locuphoric (1/2)	allophoric (3)	total
A	803	202	1005
P	47	958	1005

### Animacy of R and T

Tahitian (Snyder 2003: 80, 329)

	animate	inanimate	total
R	329	0	329
T	11	318	329

English (Thompson 1990: 243)

	animate	inanimate	total
R	190	6	196
T	9	187	196

### Full nominality of R and T

English (Thompson 1990: 244)

	person form	full nominal	total
R	115	81	196
T	11	185	196

### Givenness of R and T

Finnish (Kaiser 2002: 9, 81)

	old	new	total
R	58	23	81
T	19	62	81

### Person of R and T

German (Haspelmath 2004: 35)

	locuphoric (1/2)		allophoric (3)	total
R	147	94		241
T	49	192		241

## References

- Aikhenvald, Alexandra. 2010. *Imperatives and commands*. Oxford: Oxford University Press.
- Aristar, Anthony Rodrigues. 1997. Marking and hierarchy types and the grammaticalization of case-markers. *Studies in Language* 21(2). 313–368. doi:[10.1075/sl.21.2.04ari](https://doi.org/10.1075/sl.21.2.04ari).
- Becher, Jutta. 2005. Ditransitive Verben und ihre Objekte im Wolof: Positionsregeln und Kombinierbarkeit. *Hamburger afrikanistische Arbeitspapiere (HAAP)* 3. 13–27.
- Bossong, Georg. 1985. *Differenzielle Objektmarkierung in den neuiranischen Sprachen*. Tübingen: Narr.
- Bossong, Georg. 1991. Differential object marking in Romance and beyond. In Douglas Kibbee & Dieter Wanner (eds.), *New analyses in Romance linguistics*, 143–170. Amsterdam: Benjamins.
- Coghill, Eleanor. 2010. Ditransitive constructions in the Neo-Aramaic dialect of Telkepe. In Andrej L. Malchukov, Martin Haspelmath & Bernard Comrie (eds.), *Studies in ditransitive constructions: A comparative handbook*, 221–242. Berlin: De Gruyter Mouton.
- Creissels, Denis & Céline Mounole. 2011. Animacy and spatial cases: Typological tendencies, and the case of Basque. In Seppo Kittilä, Katja Västi & Jussi Ylikoski (eds.), *Case, animacy and semantic roles*, 155–182. Amsterdam: Benjamins.
- Croft, William. 1991. *Syntactic categories and grammatical relations: The cognitive organization of information*. Chicago: University of Chicago Press.
- D'Alessandro, Roberta. 2017. When you have too many features: Auxiliaries, agreement and clitics in Italian varieties. *Glossa: A journal of general linguistics* 2(1). 1–36. doi:[10.5334/gjgl.102](https://doi.org/10.5334/gjgl.102).
- Dahl, Östen & Kari Fraurud. 1996. Animacy in grammar and discourse. In Thorstein Fretheim & Jeanette K. Gundel (eds.), *Reference and referent accessibility*, 47–64. Amsterdam: Benjamins.
- Dalrymple, Mary & Irina Nikolaeva. 2011. *Objects and information structure*. Cambridge: Cambridge University Press.
- Dixon, R.M.W. 1994. *Ergativity*. Cambridge: Cambridge University Press.
- Geniušienė, Emma. 1987. *The typology of reflexives*. Berlin: Mouton de Gruyter.
- Grimm, Scott. 2018. Grammatical number and the scale of individuation. *Language* 94(3). 527–574.
- Haig, Geoffrey & Stefan Schnell. 2016. The discourse basis of ergativity revisited. *Language* 92(3). 591–618. doi:[10.1353/lan.2016.0049](https://doi.org/10.1353/lan.2016.0049).
- Haiman, John. 1983. Iconic and economic motivation. *Language* 59(4). 781–819. doi:[10.2307/413373](https://doi.org/10.2307/413373).

- Haspelmath, Martin & Andres Karjus. 2017. Explaining asymmetries in number marking: Singulatives, pluratives, and usage frequency. *Linguistics* 55(6). 1213–1235. doi:[10.1515/ling-2017-0026](https://doi.org/10.1515/ling-2017-0026).
- Haspelmath, Martin. 1999. Explaining article-possessor complementarity: Economic motivation in noun phrase syntax. *Language* 75(2). 227–243. doi:[10.2307/417260](https://doi.org/10.2307/417260).
- Haspelmath, Martin. 2004. Explaining the Ditransitive Person-Role Constraint: A usage-based approach. *Constructions* 2. doi:[10.5281/zenodo.831408](https://doi.org/10.5281/zenodo.831408).  
<https://journals.linguisticsociety.org/elanguage/constructions/article/view/3073.html>.
- Haspelmath, Martin. 2008. A frequentist explanation of some universals of reflexive marking. *Linguistic Discovery* 6(1). 40–63. doi:[10.1349/PS1.1537-0852.A.331](https://doi.org/10.1349/PS1.1537-0852.A.331).
- Haspelmath, Martin. 2016. Universals of causative and anticausative verb formation and the spontaneity scale. *Lingua Posnaniensis* 58(2). 33–63. doi:[10.1515/linpo-2016-0009](https://doi.org/10.1515/linpo-2016-0009)
- Haspelmath, Martin. 2017. Explaining alienability contrasts in adpossession constructions: Predictability vs. iconicity. *Zeitschrift für Sprachwissenschaft* 36(2). 193–231. doi:[10.1515/zfs-2017-0009](https://doi.org/10.1515/zfs-2017-0009).
- Haspelmath, Martin. 2019c. Differential place marking and differential object marking. *STUF – Language Typology and Universals*. (to appear)
- Haspelmath, Martin. 2020a. Explaining grammatical coding asymmetries: Form-frequency correspondences and predictability. *to appear*. <https://ling.auf.net/lingbuzz/004531>.
- Haspelmath, Martin. 2020b. Role-reference associations and the explanation of argument coding splits. *to appear*. <https://ling.auf.net/lingbuzz/004047>.
- Iemmolo, Giorgio. 2014. Symmetric and asymmetric alternations in direct object encoding. *STUF – Language Typology and Universals* 66(4). 378–403. doi:[10.1524/stuf.2013.0019](https://doi.org/10.1524/stuf.2013.0019).
- Keine, Stefan. 2010. *Case and agreement from fringe to core: A minimalist approach*. Berlin: De Gruyter.
- Maslova, Elena. 2003. *A grammar of Kolyma Yukaghir* (Mouton Grammar Library 27). Berlin: Mouton de Gruyter.
- Rude, Noel. 2009. Transitivity in Sahaptin. *Northwest Journal of Linguistics* 3(3). 1–37.
- Siewierska, Anna. 2010. Person asymmetries in zero expression and grammatical function. In Franck Floricic (ed.), *Essais de typologie et de linguistique générale: Mélanges offerts à Denis Creissels*, 471–485. Lyon: ENS Éditions.
- Snyder, Kieran. 2003. *The relationship between form and function in ditransitive constructions*. Philadelphia: University of Pennsylvania Ph.D. dissertation.
- Thompson, Sandra A. 1990. Information flow and “Dative Shift” in English. In Jerrold Edmondson, Katherine Feagin & Peter Mühlhäusler (eds.), *Development and diversity: Linguistic variation across time and space*, 239–252. Dallas: Summer Institute of Linguistics.