

Curso extracurricular:
**Comparing grammars worldwide:
Verbal and nonverbal clause constructions**
(UNAM, 20-23 October 2025)

MARTIN HASPELMATH

1. Clause constructions (Oct 20):

- A. Marking clausal arguments
- B. Issues of methodology
- C. (I) Functions vs. structures and (II) efficiency explanations

2. Alignment and differential coding (Oct 21):

- A. Accusative and ergative alignment
- B. Argument flagging splits
- C. Argument indexing vs. argument flagging

3. Nonverbal clause constructions (Oct 22):

- A. Overview
- B. Existential constructions
- C. Adjectives vs. verbs vs. nouns

4. Voice constructions (Oct 23):

- A. Causatives and anticausatives
- B. Passives
- C. Antipassives and inverses

Clause constructions:

1A. Marking clausal arguments

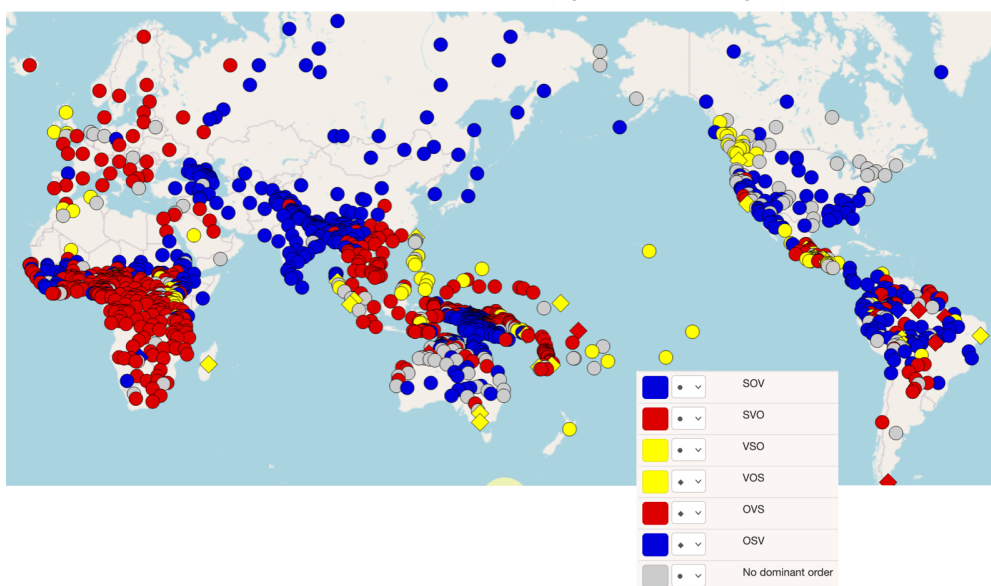
MARTIN HASPELMATH
Max Planck Institute for Evolutionary Anthropology

1. Word order of clausal arguments

Since Greenberg (1963), it has been well-known that the order of the subject and the object with respect to the verb is an important parameter of cross-linguistic variation.

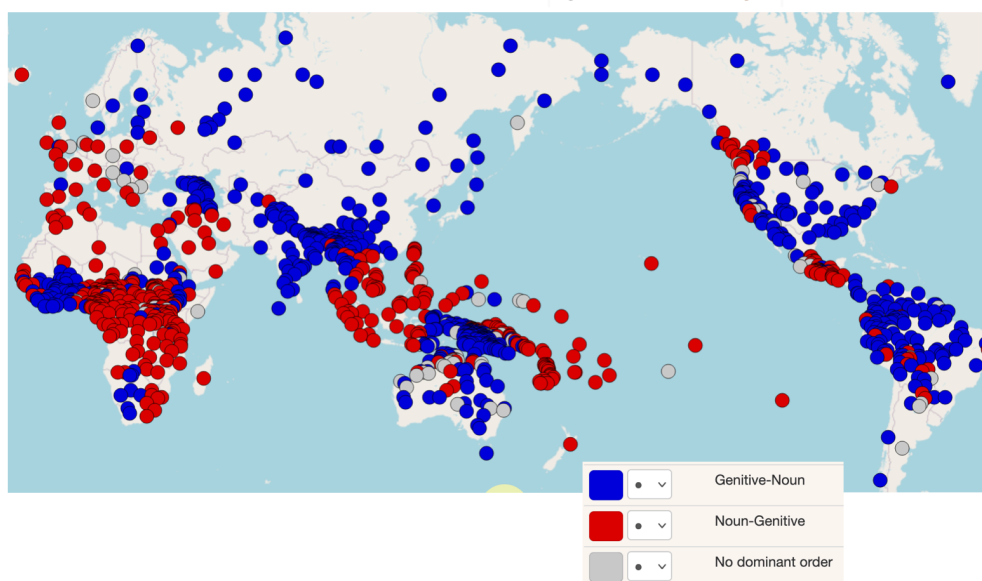
Order of Subject, Object and Verb

by Matthew S. Dryer



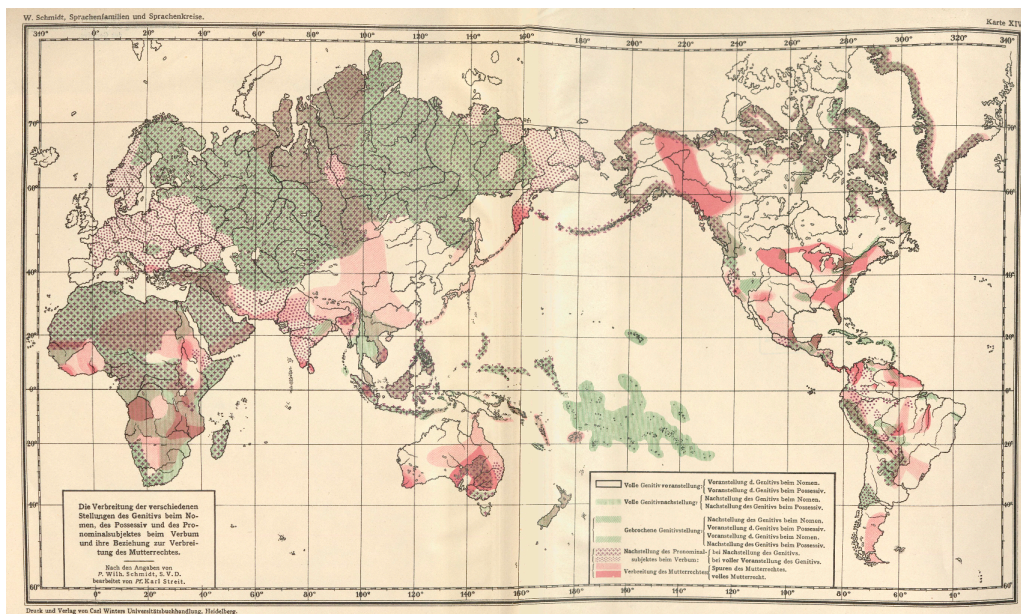
Order of Genitive and Noun

by Matthew S. Dryer



We published the *World atlas of language structures* in 2005, but interest in worldwide typological comparison is older.

As early as 1926, **Schmidt (1926)** published a work on the world's languages that contains some world maps of structural features, e.g.



But for a long time, linguists thought that the most important typological parameter has to do with the internal structure of words:

- isolating languages
- agglutinating languages
- fusional languages
- incorporating languages, etc.

This way of doing “morphological typology” goes back to German Romantics of the 19th century, e.g. **Wilhelm von Humboldt (1836)**.

Einverleibungssystem der Sprachen.

ist, die regierenden und regierten Theile des Satzes an dasselbe an, und giebt dieser Verknüpfung durch Lautformung das Gepräge eines verbundenen Ganzen: *ni-naca-qua*, ich esse Fleisch. Man könnte diese Verbindung des Substantivs mit dem Verbum als ein zusammengesetztes Verbum, gleich dem Griechischen *κρεωφαγέω*, ansehen; die Sprache nimmt es aber offenbar anders. Denn wenn aus irgend einem Grunde das Substantivum nicht selbst einverleibt wird, so ersetzt sie es durch das Pronomen der dritten Person, zum deutlichen Beweise, daß sie mit dem Verbum, und in ihm enthalten, zugleich das Schema der Construction zu haben verlangt: *ni-ē-qua in nacatl*, ich esse es. das Fleisch. Der Satz soll, seiner Form nach.



Wilhelm von Humboldt
(1767-1835)

Classical Nahuatl:

ni-naca-qua
I-meat-eat

(*ni-naka-kwa*)

(“incorporating”)

ni-c-qua in nacatl
I-it-eat the meat

However, “morphological typology” was based on too few data and too much speculation – I criticized it in Haspelmath (2009).

Moreover, “morphological typology” presupposes that we can **identify “words”** – but this is not necessarily the case (see Haspelmath 2011).

Maybe “words” exist primarily in the spelling of those languages that have spaces between “words”?

Cf. this metro station in Beijing (安华桥, *an-hua-qiao*):



Anhua Qiao



Anhuaqiao



An Hua Qiao

Constituent order typology, by contrast, is universally applicable

– instead of saying, e.g., that “X is an agglutinating language”, it’s more informative to say “X is an SVO language”.

2. Argument marking by grammatical markers

In addition to the order of constituents, basic grammatical patterns often include **argument marking**.

Traditionally: “the subject is in nominative case, the (direct) object in accusative case”

(1) Russian

a. *Sergej-Ø ljubi-t Nin-u.*
Sergey-NOM love-3SG Nina-ACC
‘Sergey loves Nina.’

b. *Nin-a ljubi-t Sergej-a.*
Nina-NOM love-3SG Sergey-ACC
‘Nina loves Sergey.’

Traditionally: “the verb agrees with the subject in person and number”

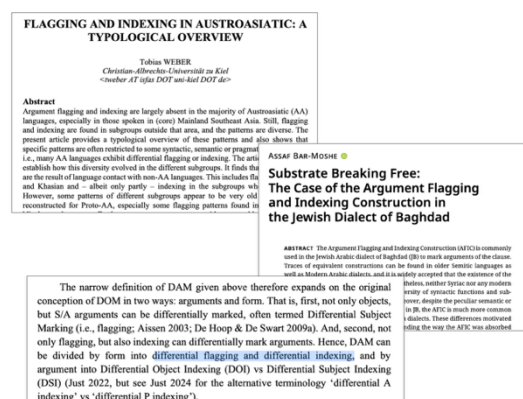
(2) Italian

a. *Lo studente lavor-a qui.*
the student work-3SG here
‘The student works here.’

- b. *Gli student-i lavor-**ano** qui.*
 the student-PL work-3PL here
 ‘The students work here.’

More recent terminology in typology:

flagging and indexing



flags: case-markers or adpositional markers (often hard to distinguish)

- (3) Japanese

Seikoo -ga tomodati -ni manga -o age-ta.
 Seiko -NOM friend -DAT manga -ACC give-PST
 ‘Seiko gave her friend the manga.’

- (4) French

Seiko donna le manga à- son ami.
 Seiko gave the manga DAT- her friend
 ‘Seiko gave her friend the manga.’

(**flagging** here replaces “case-marking or adpositional marking”)

indexes: verbal person (and number) markers, most often not literally “agreeing”

- (5) Italian *lavor-o* ‘I work’ (subject indexing)
lavor-i ‘you work’
lavor-a ‘she works’
- (6) Maltese *in-hobb-ok* ‘I love you’ (subject and object indexing)
in-hobb-u ‘I love him’
t-hobb-ni ‘you love me’

(**indexing** (= bound person marking) here replaces “agreement or cross-referencing”)

That flagging and indexing are complementary ways of marking arguments has been discussed by typologists since the 1980s (Lehmann 1982; 1985), but usually, they have used the terminology

“case and agreement”

(e.g. Siewierska 1997; Siewierska & Bakker 2009; Keine 2010; Baker 2013)

The “**flagging and indexing**” terminology has become more widespread due to my work since 2005 (Haspelmath 2005; 2013; 2019).

3. Head marking and dependent marking (Nichols 1986)

A closely related contrast is that of **dependent marking** (very similar to *flagging*) and **head marking** (similar to indexing).

A recent formulation (Lander & Nichols 2020):

“Head/dependent marking typologizes word and phrase morphosyntax, constructions, and whole languages according to whether grammatical relations are overtly marked on the head of the relevant phrase, a non-head, both, or neither. ...

- (i) a. Hungarian
az ember ház-a
 the man house-3SG.POSS
 ‘the man’s house’
- b. English
the man’s house

The essential theoretical point is that the syntax of the Hungarian and English phrases is identical (at least to the extent that it is comparable) but the morphology is opposite.”

The notions of “**head**” and “**dependent**” became well-known only after Nichols (1986) – in earlier times, linguists talked about “subjects”, “objects” and “modifiers”, but there was no general notion of *dependent*.

Before Jackendoff (1977), there was no widely-known notion of *head* either. So these notions are fairly recent, and not as stable as the notion of word order.

Nichols (1986; 1992) claimed that entire languages can be typologized as “**dependent-marking**” (e.g. Tagalog) vs. “**head-marking**” (e.g. Yucatec Maya).

(7) Tagalog (Filipino)

- a. *Hinabol nang+aso ang+pusa.*
 chased ERG+dog ABS+cat
 ‘The dog chased the cat.’
- b. *aso nang+babae*
 dog gen+girl
 ‘the girl’s dog’

(8) Yucatec Maya

- a. *T-u=nes-ah-Ø hun-tíul pàal.*
 PFV-3SG.ERG=grow-CMPL-3SG.ABS one-CLF child
 ‘The cockroach bit a child.’ (Bohnenmeyer et al. 2016: 177)
- b. *u=k’àaba’ le xch’up-pàal-a’*
 3SG.POSS=name the woman-child-DEM1
 ‘this girl’s name’ (Lehmann 2002: 42)

Nichols (1986; 1992), Nichols & Bickel (2005):
languages tend to be either generally dependent-marking or generally head-marking.

However, there are many mixed cases, e.g.

Spanish *el perro **de**+la niña* (dependent-marking)

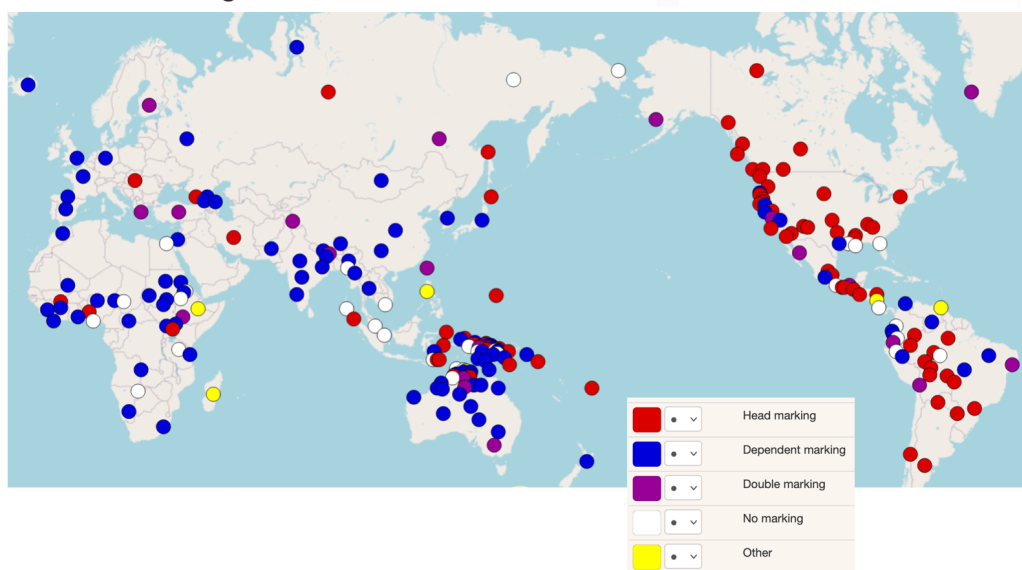
‘the girl’s dog’

*nosotros **lo**+vi-**mos** **a**+él* (head- and dependent-marking)

‘WE saw HIM’

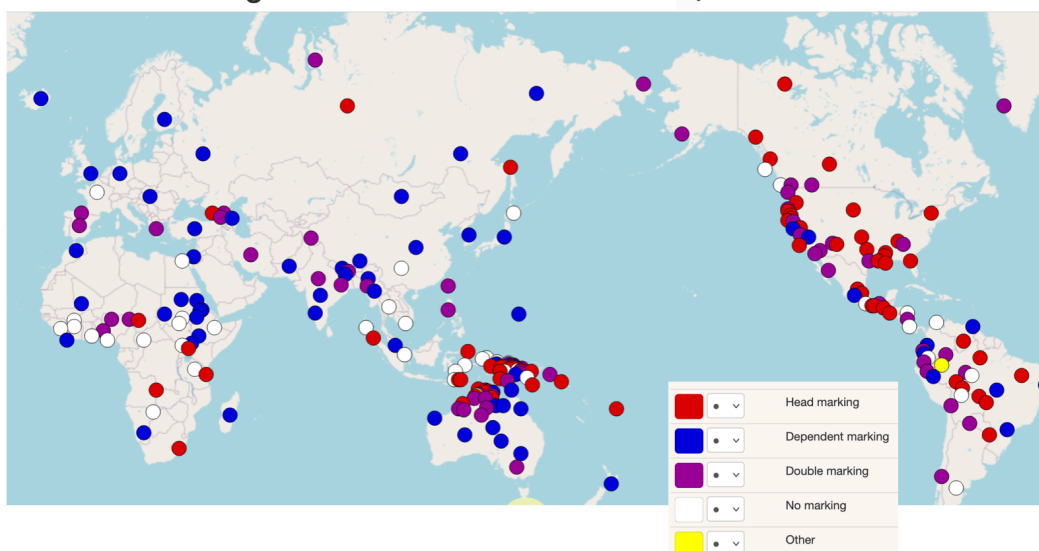
Locus of Marking in Possessive Noun Phrases

by Johanna Nichols and Balthasar Bickel

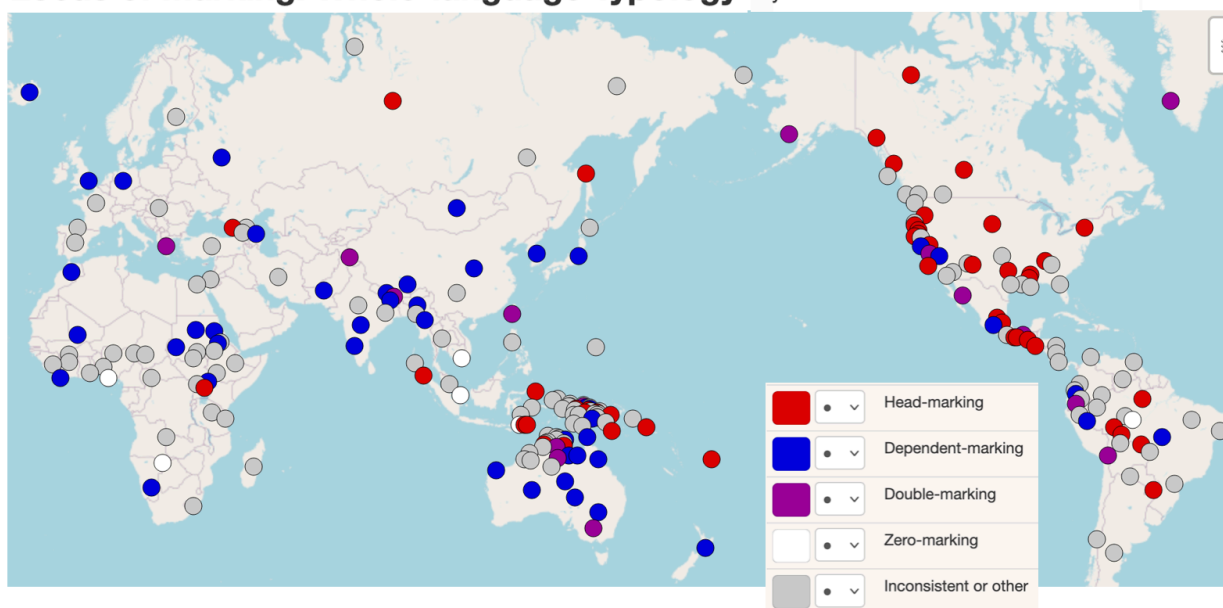


Locus of Marking in the Clause

by Johanna Nichols and Balthasar Bickel



Locus of Marking: Whole-language Typology by Johanna Nichols and Balthasar Bickel



Probably a better universal (cf. Lehmann 1982):

The further an element is to the left of the scale, the more likely it is to be indexed; and the further an element is to the right, the more likely it is to be flagged.

The involvement scale

subject > patient object > other object > beneficiary/instrument > place

←. indexing

flagging →

Yucatec (a “head-marking language”) has flagging with recipient objects:

le xch'úupal-o' t-u tsikbal-t-ah hun-p'éel che'h-bil tsikbal
DEM girl-D2 PRFV-SBJ.3 talk-TRR-CMPL one-CL.INAN laugh-GER story

ti' le xibpal-o'
LOC DEM boy-D2
'the girl told the boy a funny story'

English (a “dependent-marking language”) has indexing with subjects:

*The girl tell-**s** a story to the boy.*

4. Why *indexing* is better than “head-marking” or “agreement”

4.1. *Indexing* is better than “head-marking”

– “Head-marking” presupposes a notion of “head”, which is not as straightforward as Nichols (1986) thought

(in generative syntax since the 1990s, most “heads” are “functional heads”, and these are quite different animals)

– head-marking need not be person-marking – it can be “construct-marking”, e.g.

- (9) a. Hebrew
tmuna-t ha-yalda (cf. *tmuna* ‘picture’)
 picture-CSTR DEF-girl
 ‘the girl’s picture’
- b. Mende
ndopó-i l-oko-i (cf. *tokó* ‘arm’)
 child-DEF CSTR-arm-DEF
 ‘the child’s arm’ (Creissels 2009a: 80)

Construction markers which occur on the “head” but are not person markers are not common, and they have no particular similarity to person markers.

– indexes have a “pointing function” – they are often thought of as “referential devices”

Spanish *Cre-o que tien-es razón.*
 ‘**I** think that **you** are right.’

4.2. *Indexing* is better than “agreement”

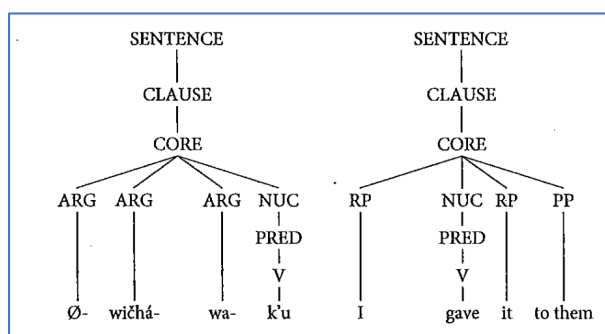
– With first and second person forms, it is odd to say that they “agree with” elements that are not present “on the surface”

underlying *Yo cre-o que tú tien-es razón.*
 surface – *creo que – tienes razón.*

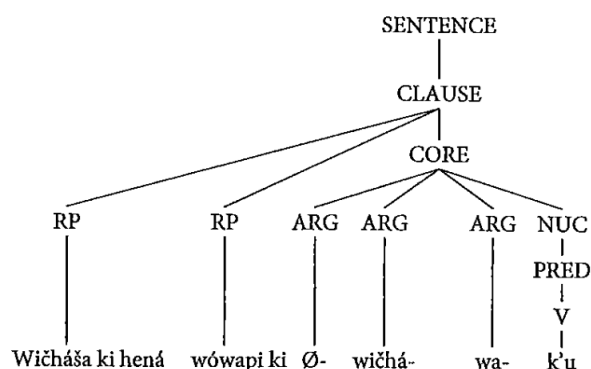
The idea that there is “**pro-drop**” here is quite strange – and apparently based on German or English.

A better way of conceptualizing this is by saying that the person indexes are the arguments, and that the additional pronouns are **conominals** (Haspelmath 2013).

This is similar to what Van Valin (2013) says:



In Van Valin's system, the conominals are outside the CORE, but inside the CLAUSE:



This is better than the traditional system, but person indexes often show variable behaviour, and it is often unclear under what precise conditions a person index occurs.

Spanish	<i>Lo+vimos a+él.</i>	'We saw him.'
	<i>?Lo+vimos al padre.</i>	'We saw the father.'

It seems that we can tell more readily whether an element is a person index than whether it "is the true argument".

5. How widespread is transitive encoding?

(cf. Haspelmath 2015)

Languages differ in the extent to which they make use of transitive encoding, as is clear from English/German contrasts (cf. Hawkins 1986):

in German, verbs like 'help' and 'follow' are not encoded transitively, unlike English:

- | | | |
|--------|---------|--|
| (10)a. | English | <i>He_{NOM} helped her_{ACC}.</i> |
| b. | German | <i>Er_{NOM} half ihr_{DAT}.</i> |
| (11)a. | English | <i>They_{NOM} followed them_{ACC}.</i> |
| b. | German | <i>Sie_{NOM} folgten ihnen_{DAT}.</i> |

In other languages, the encoding is more transitive than in English:

(12) Bora (Boran; Columbia)

wajpi ihjyúcunú ováhtsa-ke
man shout boy-ACC

‘The man shouted **at** the boy.’ (Seifart 2013)

(13) Yaqui (Uto-Aztecan; Mexico)

Inepo chu'u-ta jariwa.
1sg.SBJ dog-ACC look.for

‘I am looking **for** my dog.’ (Estrada et al. 2013)

– Can we predict transitive encoding from the meaning of a verb? (cf. Tsunoda 1985; Malchukov 2005)

– Are there implicational universals such as (14)?

(14) If a language has transitive encoding for ‘like’, it will also have transitive encoding for ‘see’

First task: **descriptive** information on many different languages

our solution: the **ValPaL database** (valpal.info, Hartmann et al. 2013)

Second task: **comparative concepts** for comparison: (cf. Haspelmath 2010)
how to recognize a transitive clause in different languages?

– transitive clause = a clause with subject and object?

(but what are “subject” and “object”?
grammatical relations are not universal, cf. Dryer 1997)

(12) English: *The car **weighs** two tons.*

(13) Icelandic (Barðdal 2013)

Mig fýsir þess.
I-ACC long that.GEN

‘I long for that.’

(14) Lezgian (Nakh-Daghestanian, Russia; Haspelmath 1993)

Ali-di kic'-Ø q'e-na.
Ali-ERG dog-ABS kill-PST

‘Ali killed the dog.’

our solution:

– a verb is transitive if it has two arguments that are coded like the two arguments of ‘break’ (the **A-argument** and the **P-argument**; Haspelmath 2011)

– thus, “transitive” is a **mixed functional-formal** comparative concept: it is based on the **functional** concept ‘break’, and the **formal** concept ‘identity of argument coding’

Third task: a (**universal**) ranking of verbs by “transitivity prominence” (Haspelmath 2015),
i.e. the proportion of verbs with an A-argument and a P-argument

(Fourth task: **Explanation:** metaphorical extension of the ‘break’ pattern to other verbs? cf. Croft 1991)

ValPaL

Home

Languages

Verb meanings

All coding frames

Microroles

All alternations

Languages

Showing 1 to 36 of 36 entries

← Previous

1

Next →

1

Name	Glottocode	Family	Latitude	Longitude	Contributors
<input type="text" value="Search"/>	<input type="text" value="Search"/>	--any--			
Ainu (Southern Hokkaido)	ainu1240	Ainu	43	143	Anna Bugaeva
Balinese	bali1278	Austronesian	-8.36	115.08	Masayoshi Shibatani and Ketut Artawa
Bezhta	bezh1248	Nakh-Daghestanian	42.12	46.03	Zaira Khalilova and Bernard Comrie
Bora	bora1263	Boran	-2.16666666667	-72.3333333333	Frank Seifart
Chintang	chhi1245	Sino-Tibetan	26.947628	87.211189	Robert Schikowski and Balthasar Bickel and Netra Prasād Paudyal
Eastern Armenian (standard Eastern Armenian)	east2283	isolate	40	45	Victoria Khurshudian and Michael Daniel
Emai	emai1241	Atlantic-Congo	7.08333333333	5.91666666667	Ronald Schaefer and Francis Egbokhare
English	stan1293	Indo-European	53	-1	Cliff Goddard
Even	even1260	Tungusic	70.67	130.91	Andrej Malchukov
Evenki	even1259	Tungusic	56	125	Igor Nedjalkov
German (Standard)	stan1295	Indo-European	51.344339	12.348633	Martin Haspelmath and Luisa Baumann
Hokkaido Japanese	hokk1249	Japonic	43.691708	142.866211	Kan Sasaki

Table 2. The 80 verb meanings of ValPaL

EAT	MEET	PEEL	RUN	ROLL (intr)
HUG	TALK	HIDE	SIT	SINK (intr)
LOOK AT	ASK FOR	SHOW	SIT DOWN	BURN (intr)
SEE	SHOUT AT	GIVE	JUMP	BE DRY
SMELL	TELL	SEND	SING	RAIN
FEAR	SAY	CARRY	GO	BE A HUNTER
FRIGHTEN	NAME	THROW	LEAVE	GRIND
LIKE	BUILD	TIE	LIVE	WIPE
KNOW	BREAK	PUT	LAUGH	DIG
THINK	KILL	POUR	SCREAM	PUSH
SEARCH FOR	BEAT	COVER	FEEL PAIN	BRING
WASH	HIT	FILL	FEEL COLD	STEAL
DRESS (tr)	TOUCH	LOAD	DIE	TEACH
SHAVE	CUT	BLINK	PLAY	HEAR
HELP	TAKE	COUGH	BE SAD	COOK
FOLLOW	TEAR	CLIMB	BE HUNGRY	BOIL (intr)

The valency frames give information both about **flagging** and **indexing**:

(15) Bora (Seifart 2013)

a. COVER	<i>wátájcó</i>	<1-nom 2-acc 3-all V>
coverer	1	NP-nom
covered thing	2	NP-acc
cover	3	NP-all
b. FOLLOW	<i>úraavyé</i>	<1-nom 2-acc>
follower	1	NP-nom
followee	2	NP-acc
c. ROLL (intr)	<i>víyiivye</i>	<1-nom 2-all V>
rolling entity	1	NP-nom
rolling location	2	NP-all

(16) Ewen (Malchukov 2013)

a. BREAK	<i>čelgel-</i>	<1-nom 2-acc 3-instr V. subj [1]>
breaker	1	1-nom & V.subj
broken thing	2	2-acc
breaking instrument	3	3-instr
b. HELP	<i>bele-</i>	<1-nom 2-dat V. subj [1]>
helper	1	1-nom & V.subj
helpee	2	2-dat

(17) Hoocək (Hartmann 2013)

a. BREAK	<i>gišiš</i>	<1 2 und [2]. act [1].V>
breaker	1	act.V
broken thing	2	und.V
b. LOOK AT	<i>horoğoc</i>	<1 2 und [2]. act [1].V>
looker	1	act.V
looked at entity	2	und.V

Differences between languages

ValPaL languages ranked by transitivity-prominence

(= percentage of transitively encoded verbs among the sample verbs)

Chintang	.77	Evenki	.63
Emai	.73	Hokkaido Japanese	.63
Ojibwe	.71	Modern Standard Arabic	.62
Yucatec Maya	.69	Jaminjung	.62
Zenzontepec Chatino	.68	Xârâcùù	.61
Yaqui	.68	English	.60
Balinese	.68	German	.60
Bora	.67	Ainu	.57
Jakarta Indonesian	.67	Russian	.56
Even	.65	Nen	.56
Japanese (standard)	.65	Eastern Armenian	.55
Sliammon	.65	Yorùbá	.54
Mapudungun	.65	Sri Lanka Malay	.53
Hoocək	.64	Ket	.52
Njuu	.64	Mandarin Chinese	.50
Mandinka	.64	Icelandic	.48
Mitsukaido Japanese	.64	Bezhta	.46
Italian	.64		

Some examples of unexpected transitive encoding in non-European languages:

- (10) Chintang (Schikowski et al. 2013)

maʔmi-ŋa bħiri sopt-o-s-e
 person-ERG hill climb-[3sA.]3[s]P-PRF-IND.PST
 ‘The man has climbed (up) the hill.’

- (11) Emai (Schaefer & Egbokhare 2013)

Òjè ò ó jè àlèkè.
 Oje SC C laugh Aleke
 ‘Oje is laughing at Aleke.’

- (12) Yucatec Maya (Lehmann 2013)

le xibpal-o' t-u pakat-ah le xch'úupal-o'
 DEM boy-D2 PRFV-SBJ.3 look-CMPL DEM girl-D2
 ‘The boy looked at the girl.’

- (13) Yaqui (Estrada-Fernández et al. 2013)

Ume yoemem ili uusita jariwa.
 DET.PL man-PL little child-ACC search_for
 ‘The men are looking for the child.’

- (14) Balinese (Shibatani & Artawa 2013)

Tiang nyakitang lima=n tiang=e.
 I hurt hand=LIG I=POSS
 ‘I am feeling pain in my arms.’ (Lit. ‘I am hurting my arms.’)

Differences between verbs

Tsunoda (1985) had proposed the following scale of transitive encoding:

- (15) direct effect > perception > pursuit > cognition > emotion
 BREAK > HIT > SEE > LOOK SEARCH KNOW LIKE

- (16) transitivity-prominence (ValPaL-based)

BREAK	1.00
HIT	.94
SEE	.92
LOOK AT	.74
SEARCH	.89
KNOW	.88
LIKE	.77

ValPaL verb meanings ranked by transitivity-prominence

(= percentage of transitively encoded verbs among all counterpart verbs)

BREAK	1.00	BRING	.90	LEAVE	.44
FRIGHTEN	1.00	BUILD	.90	SHOUT AT	.43
HIDE	1.00	COVER	.90	SING	.36
KILL	.98	SEARCH FOR	.89	FEEL PAIN	.12
TEAR	.97	HUG	.89	LAUGH	.11
STEAL	.97	KNOW	.88	PLAY	.10
CUT	.97	COOK	.88	BLINK	.08
TAKE	.96	LOAD	.87	RUN	.08
FILL	.96	POUR	.86	BE SAD	.07
WASH	.96	SEND	.85	SIT	.06
PEEL	.96	HEAR	.84	LIVE	.05
WIPE	.95	SMELL	.83	COUGH	.03
BEAT	.95	SHAVE	.81	SIT DOWN	.03
PUSH	.95	TOUCH	.80	SCREAM	.03
HIT	.94	DIG	.80	SINK (intr.)	.03
GIVE	.93	TELL	.78	BOIL (intr.)	.03
CARRY	.93	NAME	.78	JUMP	.02
EAT	.93	LIKE	.77	FEEL COLD	0
THROW	.93	HELP	.77	DIE	0
PUT	.93	LOOK AT	.74	BE HUNGRY	0
ASK FOR	.92	FOLLOW	.74	BE DRY	0
DRESS (tr)	.92	MEET	.72	RAIN	0
GRIND	.92	FEAR	.55	BE A HUNTER	0
SEE	.92	THINK	.53	ROLL (intr.)	0
TEACH	.91	CLIMB	.53	BURN (intr.)	0
TIE	.91	SAY	.49		
SHOW	.90	TALK	.45		

References

- Baker, Mark C. 2013. On agreement and its relationship to case: Some generative ideas and results. *Lingua* (SI: Syntax and Cognition: Core Ideas and Results in Syntax) 130. 14–32.
(doi:[10.1016/j.lingua.2012.03.010](https://doi.org/10.1016/j.lingua.2012.03.010))
- Creissels, Denis. 2009. Construct forms of nouns in African languages. In Peter K. Austin, David Nathan & Peter Sells, Oliver Bond, Monik Charette (ed.), *Proceedings of Conference on Language Documentation and Linguistic Theory 2*, 73–82. London: SOAS.
(<http://www.deniscreissels.fr/public/Creissels-constr.afr.2009.pdf>)
- Croft, William. 1991. *Syntactic categories and grammatical relations: The cognitive organization of information*. Chicago: University of Chicago Press.
- Estrada Fernández, Zarina & Tubino, Mercedes & Villalpando, Jesús. 2013. Yaqui valency patterns. In Hartmann, Iren & Haspelmath, Martin & Taylor, Bradley (eds.), *Valency Patterns Leipzig*. Leipzig: Max Planck Institute for Evolutionary Anthropology.
(<http://valpal.info/languages/yaqui>)
- Greenberg, Joseph H. 1963. Some universals of grammar with particular reference to the order of meaningful elements. In Greenberg, Joseph H. (ed.), *Universals of language*, 73–113. Cambridge, MA: MIT Press.
- Haspelmath, Martin. 1993. *A grammar of Lezgian* (Mouton Grammar Library 9). Berlin: Mouton de Gruyter.
- Haspelmath, Martin. 2005. Argument marking in ditransitive alignment types. *Linguistic Discovery* 3(1). 1–21.

- Haspelmath, Martin. 2009. An empirical test of the Agglutination Hypothesis. In Scalise, Sergio & Magni, Elisabetta & Bisetto, Antonietta (eds.), *Universals of language today*, 13–29. Dordrecht: Springer. (Accessed May 27, 2016.)
- Haspelmath, Martin. 2011. On S, A, P, T, and R as comparative concepts for alignment typology. *Linguistic Typology* 15(3). 535–567.
- Haspelmath, Martin. 2013. Argument indexing: A conceptual framework for the syntax of bound person forms. In Bakker, Dik & Haspelmath, Martin (eds.), *Languages across boundaries: Studies in memory of Anna Siewierska*, 197–226. Berlin: De Gruyter Mouton. (<https://zenodo.org/record/1294059>)
- Haspelmath, Martin. 2015. Transitivity prominence. In Malchukov, Andrej L. & Comrie, Bernard (eds.), *Valency classes in the world's languages: A comparative handbook (Volume 1)*, 131–147. Berlin: De Gruyter Mouton. (<https://zenodo.org/deposit/4551122>)
- Haspelmath, Martin. 2019. Indexing and flagging, and head and dependent marking. *Te Reo* 62(1). 93–115. (doi:[10.17617/2.3168042](https://doi.org/10.17617/2.3168042))
- Hawkins, John A. 1986. *A comparative typology of English and German: Unifying the contrasts*. Austin: University of Texas Press.
- Humboldt, Wilhelm von. 1836. *Ueber die Verschiedenheit des menschlichen Sprachbaues und ihren Einfluss auf die geistige Entwicklung des Menschengeschlechts, von Wilhelm von Humboldt*. Berlin: Druckerei der königlichen Akademie der Wissenschaften. (<http://catalogue.bnf.fr/ark:/12148/cb30628312k>) (Accessed October 24, 2025.)
- Jackendoff, Ray. 1977. *X-bar syntax: A study of phrase structure*. Cambridge, MA: MIT Press.
- Keine, Stefan. 2010. *Case and agreement from fringe to core: A minimalist approach*. Berlin: De Gruyter.
- Lander, Yury & Nichols, Johanna. 2020. Head/dependent marking. *Oxford Research Encyclopedia of Linguistics*. (doi:[10.1093/acrefore/9780199384655.013.523](https://doi.org/10.1093/acrefore/9780199384655.013.523)) (<https://oxfordre.com/view/10.1093/acrefore/9780199384655.001.0001/acrefore-9780199384655-e-523>) (Accessed March 7, 2020.)
- Lehmann, Christian. 1982. Universal and typological aspects of agreement. In Seiler, Hansjakob & Stachowiak, Franz Josef (eds.), *Apprehension: Das sprachliche Erfassen von Gegenständen*, 201–267. Tübingen: Narr. (https://www.christianlehmann.eu/publ/lehmann_agreement.pdf)
- Lehmann, Christian. 1985. On grammatical relationality. *Folia Linguistica* 19(1–2). 67–110. (doi:[10.1515/flin.1985.19.1-2.67](https://doi.org/10.1515/flin.1985.19.1-2.67))
- Lehmann, Christian. 2002. *Possession in Yucatec Maya* (ASSidUE 10). 2nd, revised edition edn. Erfurt: Universität Erfurt. (<http://www.christianlehmann.eu/publ/ASSidUE10.pdf>)
- Malchukov, Andrej. 2005. Case pattern splits, verb types, and construction competition. In Amberber, Mengistu & de Hoop, Helen (eds.), *Competition and variation in natural languages: The case for case*, 73–117. London: Elsevier.
- Nichols, Johanna. 1986. Head-marking and dependent-marking grammar. *Language* 62(1). 56–119.
- Nichols, Johanna. 1992. *Linguistic diversity in space and time*. Chicago: The University of Chicago Press.
- Nichols, Johanna & Bickel, Balthasar. 2005a. Locus of marking in possessive noun phrases. In Haspelmath, Martin & Dryer, Matthew S. & Gil, David & Comrie, Bernard (eds.), *The world atlas of language structures*, 102–105. Oxford: Oxford University Press. (<http://wals.info/chapter/24>)
- Nichols, Johanna & Bickel, Balthasar. 2005b. Locus of marking in the clause. In Haspelmath, Martin & Dryer, Matthew S. & Gil, David & Comrie, Bernard (eds.), *The world atlas of language structures*, 98–101. Oxford: Oxford University Press. (<http://wals.info/chapter/23>)
- Nichols, Johanna & Bickel, Balthasar. 2005c. Locus of marking: Whole-language typology. In Martin Haspelmath, David Gil & Bernard Comrie, Matthew S. Dryer (ed.), *The world atlas of language structures*, 106–109. Oxford: Oxford University Press.
- Schmidt, Wilhelm. 1926. *Die Sprachfamilien und Sprachenkreise der Erde*. Heidelberg: Winter.
- Seifart, Frank. 2013. Bora valency patterns. In Hartmann, Iren & Haspelmath, Martin & Taylor, Bradley (eds.), *Valency Patterns Leipzig*. Leipzig: Max Planck Institute for Evolutionary Anthropology. (<http://valpal.info/languages/yaqui>)

Clause constructions:

1B. Issues of methodology

MARTIN HASPELMATH
Max Planck Institute for Evolutionary Anthropology

1. Two ways of approaching the study of Human Language

(A) the description-comparison approach

we **describe** many languages worldwide and **compare** them

(B) the “deep reality” approach

we look for some deep “**underlying**” **structure** that is not immediately apparent

The description-comparison approach may look a bit boring –
like collecting butterflies and displaying them in a museum.

The “deep reality” approach may look more exciting –
like chemistry and physics, where it has been shown
that the stuff that surrounds us consists of elements of atoms.

But I argue that the description-comparison approach is better suited to helping us understand the nature of Human Language.

2. The study of Human Language, and of particular languages

I assume that

- we are interested in **theoretical** linguistics (not just in applied linguistics)
- we are interested in **Human Language**, not just in particular languages (*p-languages*)
- in other words, we want to do *general linguistics* (*g-linguistics*)

but how do we solve the **general linguistics paradox**?
(Haspelmath 2021b)

We want to explore and understand the nature of Human Language,
but what we can observe directly is particular languages.

answer: – we study **universals**
[– alternatively: we study nonconventional aspects of Human Language]

When we observe a particular language, we can describe it, but it does not tell us anything about Human Language directly.

Consider words such as	<i>house</i>	дом	房子
	<i>tree</i>	дерево	树
	<i>red</i>	красный	红

they differ across languages in **random, accidental** ways – they do not tell us anything about the nature of Human Language.

The situation is basically the same with morphosyntax – languages differ in **seemingly random** ways, e.g.

<i>I have a red house.</i>	
У меня красный дом.	
我有一座红房子.	[wǒ yǒu yī zuò hóng fángzi]

English has a transitive verb that shows agreement with the possessor, while Russian has no verb and a spatial preposition (*u menja* ‘near me’), and Chinese seems to have an existential verb (有 *you* ‘there is’) (‘I, there is a red house’).

Thus, particular languages (*p-languages*) are not an immediate window to human cognition – they differ across populations, and they are learned like other aspects of human culture.

so we are confronted with the **general linguistics paradox**:
(Haspelmath 2021b)

We want to explore and understand the nature of Human Language,
but what we can observe directly is particular languages.

Cf. religious studies:

What we can observe directly is particular religions.
But how do we learn from them about Human Religion in general?

The solution: **comparing cultures**:

- | | |
|-----------------------|---------------------------------|
| – comparing religions | (comparative religious studies) |
| – comparing languages | (comparative grammar studies) |

3. The “deep reality” approach:

Languages are made from the same building blocks

Example: dependent case theory (e.g. Baker 2015)

Baker (2015) claims that grammatical case-marking patterns in the world’s languages are usually determined by dependent-case rules as in (1).

- (2) a. High case in the clause is ergative.
 b. Low case in the clause is accusative.
 c. High case in VP is dative.
 e. High case in NP is genitive.
 f. Unmarked case is nominative-absolutive.

Cls:	NP	→	ERG
VP:	NP	→	DAT
VP:	NP	→	unmarked = ABS

Ali-di za-z ktab wuga-na. (Lezgian)
 Ali-ERG I-DAT book give-AOR
 ‘Ali gave me a book.’

There is no unique way to determine what is “high” and “low”, but Baker is not worried by this:

“We should not expect to find many differences between c-command and a notion like thematic prominence, since the two notions are closely related.” (p. 81)

But elsewhere, he appeals to a movement operation in order to explain some case-marking effects. Thus, Sakha (a Turkic language) has differential object marking on definite nominals (p. 125-126):

- (22) a. Masha salamaat-y sie-te.
 Masha porridge-ACC eat-PAST.3sS
 ‘Masha ate the porridge.’
 b. Masha salamaat sie-te.
 Masha porridge eat-PAST.3sS
 ‘Masha ate porridge.’

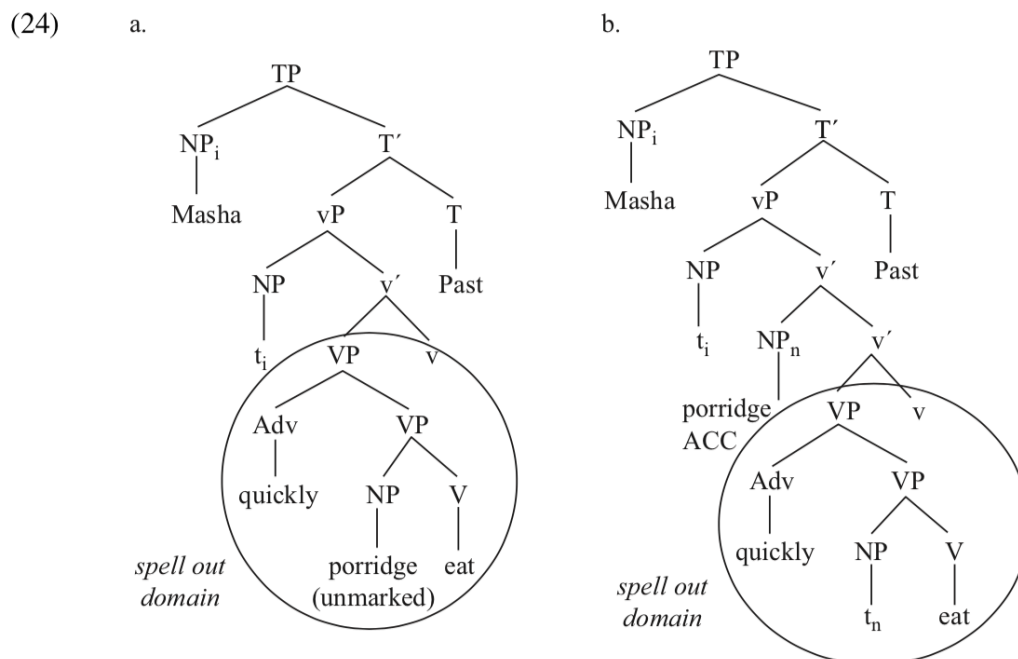
- (23) a. Masha salamaat-y türgennik sie-te.
 Masha porridge-ACC quickly eat-PAST.3sS
 ‘Masha ate the porridge quickly.’

*(Object is * without ACC)*

- b. Masha türgennik salamaat sie-te.
 Masha quickly porridge eat-PAST.3sS
 ‘Masha ate porridge quickly.’

(ACC would imply that the object is focused)

Baker's actual proposal for the movement is rather intricate and thus quite “deep” (p. 126):



Hypothesizing a “deep reality” of this kind makes sense only if the building blocks of language structure are hypothesized to be **innate**

– if they are **natural kinds**, analogous to the chemical elements (cf. Baker 2001)

In Haspelmath (2021b), I noted that there are three ways in which the study of a particular language can contribute to general linguistics (three ways to get around the “general linguistics paradox”):

(1) **nonconventional aspects of Human Language**

e.g. psycholinguistic measurements, slips of the tongue

(2) **worldwide language comparison**

e.g. Greenberg (1963), *World Atlas of Language Structures*

(3) **a natural-kinds programme for the building blocks of languages**

(i.e. the proposal that the ingredients of grammars are innate)

The **natural-kinds programme** is fascinating, and has been compared to the discovery of the chemical elements of the 19th century (Mendeleyev’s Periodic Table of Elements). The idea of an **essentialist approach** to grammar is not absurd

But we have not made much progress – we do not know what the innate categories and features of the human mind are...

4. Comparison uses comparative concepts, defined uniformly across languages

In order to compare languages systematically, one needs comparative concepts, and these need to be defined in the same way for all languages.

For example, flagging and indexing must be defined uniformly – we cannot say that a flag (case marker, adposition) is defined differently in different languages.

This principle is not always observed, as can be seen in the following quotation (from Svenonius 2007: 67):

2.2 Identifying P in different languages

Of course, there is always the hazard when comparing different languages that one is simply seeing categories that one expects to see. What if, for example, the Chinese words identified as prepositions in Section 2.1 are really verbs, or the Northern Sámi words identified in Section 2.1 as postpositions are really nouns, which have been classified as adpositions because they translate English adpositions? What evidence is there that the languages in questions have a distinct *category* P?

Such questions must be answered on a case-by-case basis, using language-internal diagnostics. Below I discuss the Chinese and Northern Sámi examples in turn, but first set the stage with a brief discussion of English.

example of Chinese:

<i>dào Shànghǎi</i>	‘to Shanghai’
	‘arrive in Shanghai’

Proposals for defining the terms *adposition*, *case marker*, *flag*, *clitic*, and *affix*:

adposition

An adposition is a clitic flag.

case marker

A case marker is an affix flag.

flag

A flag is a marker = clitic or affix) that occurs on a nominal and gives information about its semantic role or syntactic function.

clitic

A clitic is a bound morph that is not a root or an affix. (Haspelmath 2023)

affix

An affix is a bound morph that is not a root, that occurs on a root, and that cannot occur on roots of different root classes. (Haspelmath 2021a)

These definitions do not always give “intuitive” results, but our intuitions are usually based on particular phenomena we know well, not necessarily on something deep.

For example, “clitic” is widely thought to be defined phonologically, as “prosodically deficient” – however, affixes are usually “prosodically deficient” in the same way, and clitics are often prosodically complete. The stereotype does not work as a definition.

5. Comparative concepts are defined sharply, not as prototypes

Comparative concepts are **tools** for comparing languages systematically – so they must be sharp tools.

If we want to make a database, we need to be able to say whether a phenomenon falls under concept A or concept B. If A and B are not sharply defined, there is no clear answer to this.

In practice, typologists often define their terms in somewhat vague terms, e.g.

Iggesen 2005 (in WALS)

“A marker does not necessarily have to be attached to the phrasal head to be counted as nominal case; it is only required that the marker show a sufficient degree of **bondedness** (phonological integration) with its host noun ...”

But what is a “sufficient degree”?

Linguists often note that the space of variation is continuous, and therefore they feel that the terms should not be sharply defined but as “prototypes”, e.g. Fagard & Mardale (2012) about the definition of “adposition”:

“Elements which do not fit in well ... can then be described as ‘poor’ instances of the category in question: ‘pseudo’ adpositions, ‘improper’ adverbs, etc. In our view, a **prototypical framework** more adequately describes such classes: grammatical categories are then seen as heterogeneous paradigms containing elements **which display the category’s prototypical features** to various degrees. In the case of adpositions, we expect to find a core of prototypical adpositions, a set of less prototypical elements and finally a set of elements which share only some (proto)typical features of adpositions. The prototypical framework implies, besides, that **these distinctions are not clear-cut**, and that many elements can be shown to be in-between these subcategories.”

However, if a grammatical term such as *adposition* denotes a prototype rather than a sharply delineated concept,

- we cannot use it for making cross-linguistic databases
- we cannot use it to test hypotheses quantitatively

6. Defining comparative concepts is not “difficult”

Linguists often say that it is “difficult” to define a term, especially in survey articles where they are supposed to define it, e.g.

Nevis (2000: 389): “Because definitions of **clitics** are so varied and because the phenomenon itself appears to lack uniformity and cohesion, it may be difficult to provide clear, unambiguous, and theory-neutral definitions.”

Joseph (2002): “the task of defining what a **word** is in any language is a challenge”

van der Auwera & van Alsenoy (2016: 473): „the difficulty of defining **negative concord**...”

Gibson et al. (2019): „As a category, **copulas** are not easily defined. Traditional definitions tend to consider copulas to be semantically empty elements...”

Nazzi & Cutler (2019: 26): “The contrast between **consonants** and **vowels** ... is not an easy contrast to define.”

<https://dlc.hypotheses.org/2675>

However, it is not “difficult” to provide a definition for a term – definitions are often short and pithy (e.g. “A clitic is a bound morph that is not a root or an affix”, Haspelmath 2023).

What may be “difficult” is to find a “satisfactory” definition – but what is “satisfactory”? This is a subjective notion.

Linguists often note that a term is used in a variety of different ways, and in such cases, it is **impossible** (not just difficult) to provide a definition that covers all the usages in the earlier literature.

For example, Nevins’s (2000) task of surveying the literature on **clitics** is made difficult by the variety of ways in which the term *clitic* has been used. (And does it even make sense to survey “the literature on clitics” if “clitic” is used in very different ways?)

7. Switching the perspective from an essentialist approach to a measurement approach

The methodological approach that I am proposing involves comparative concepts similar to **units of measurement**, not to essentialist natural kinds (as in the “deep reality” approach).

Instead of

- looking for the “correct analysis” of a language-particular pattern
- debating the “status” of a phenomenon
- proposing “underlying” entities that allow a more elegant description

... I think we should “**measure**” the differences between languages and use these measurements to find general pattern.

Units of measurements are not claimed to correspond to anything in nature – they merely serve as yardsticks to allow scientists to compare related phenomena and to state possible generalizations.

Science should “**cut nature at its joints**”, but when nature appears to be **continuous**, a measurement approach may be fruitful.

In comparative linguistics, we use **comparative concepts** to compare languages (Haspelmath 2010).

They are **artificial concepts** designed for the purpose of “measuring” (= identifying and assessing) differences and similarities between languages.

They are not claimed to correspond to “nature’s joints”, but merely serve the practical purpose of allowing comparison, just like the symbols of the IPA.

Mendeleev vs. Passy:

The periodic table of elements shows the natural kinds of chemistry (“nature’s joints”).

But the IPA table (developed by Paul Passy and his colleagues since the 1880s) does not show the natural kinds of phonology.

It seems that we cannot identify the **natural kinds** of phonology and syntax, if there are any (at least not yet, and we do not seem to be moving closer).

The situation would then be as in **comparative anthropology**, where we cannot identify natural kinds either (maybe with the exception of kinship):

categories like “tribe”, “chief”, “trade”, “land ownership”, “religion”, “god”, “art”, “taboo” are highly variable across cultures, so in order to compare societies, anthropologists must work with **artificial comparative concepts** such as “moralizing high god” (e.g. Botero et al. 2014).

References

- Baker, Mark C. 2015. *Case: Its principles and parameters*. Cambridge: Cambridge University Press.
- Botero, Carlos A. & Gardner, Beth & Kirby, Kathryn R. & Bulbulia, Joseph & Gavin, Michael C. & Gray, Russell D. 2014. The ecology of religious beliefs. *Proceedings of*

- the National Academy of Sciences* 111(47). 16784–16789.
(doi:[10.1073/pnas.1408701111](https://doi.org/10.1073/pnas.1408701111))
- Fagard, Benjamin & Mardale, Alexandru. 2012. The pace of grammaticalization and the evolution of prepositional systems: Data from Romance. *Folia linguistica* 46(2). 303–340.
- Greenberg, Joseph H. 1963. Some universals of grammar with particular reference to the order of meaningful elements. In Greenberg, Joseph H. (ed.), *Universals of language*, 73–113. Cambridge, MA: MIT Press.
- Haspelmath, Martin. 2010. Comparative concepts and descriptive categories in crosslinguistic studies. *Language* 86(3). 663–687. (doi:[10.1353/lan.2010.0021](https://doi.org/10.1353/lan.2010.0021))
- Haspelmath, Martin. 2021a. Bound forms, welded forms, and affixes: Basic concepts for morphological comparison. *Voprosy Jazykoznanija* 2021(1). 7–28. (doi:[10.31857/0373-658X.2021.1.7-28](https://doi.org/10.31857/0373-658X.2021.1.7-28))
- Haspelmath, Martin. 2021b. General linguistics must be based on universals (or nonconventional aspects of language). *Theoretical Linguistics* 47(1–2). 1–31. (doi:[10.1515/tl-2021-2002](https://doi.org/10.1515/tl-2021-2002))
- Haspelmath, Martin. 2023. Types of clitics in the world’s languages. *Linguistic Typology at the Crossroads* 3(2). 1–59. (doi:[10.6092/issn.2785-0943/16057](https://doi.org/10.6092/issn.2785-0943/16057))
- Iggesen, Oliver A. 2005. Number of cases. In Haspelmath, Martin & Dryer, Matthew S. & Gil, David & Comrie, Bernard (eds.), *The world atlas of language structures*, 202–205. Oxford: Oxford University Press. (<http://wals.info/chapter/49>)
- Nevis, Joel A. 2000. Clitics. In Booij, Geert E. & Lehmann, Christian & Mugdan, Joachim (eds.), *Morphology: An international handbook on inflection and word-formation (Volume 1)*, 388–404. Berlin: de Gruyter. (<https://doi.org/10.1515/9783110111286.1.5.360>)
- Svenonius, Peter. 2007. Adpositions, particles and the arguments they introduce. In Reuland, Eric & Bhattacharya, Tanmoy & Spathas, Giorgos (eds.), *Argument structure*, 63–103. Amsterdam: Benjamins.

Clause constructions:

1C. (I) Functions vs. structures and (II) efficiency explanations

MARTIN HASPELMATH
Max Planck Institute for Evolutionary Anthropology

I. Functions vs. structures

The term *construction* has become quite prestigious in recent decades, especially in connection with “Construction Grammar”.

But what exactly is a *construction*?

(See Haspelmath 2023b for a general discussion of the term and a definition.)

In comparative contexts, the term “construction” is used both for notional entities and for formal entities, e.g.

notional (“what is expressed?”)

comparative construction
inalienable possession construction
experiential construction
polar question construction

formal (“how is it expressed?”)

serial verb construction
incorporation construction
ergative construction
extraposition construction

These are called *construction-functions* and *construction-strategies* here, following Croft (2022).

(Croft says “constructions” and “strategies”, but strategies are of course constructions, too.)

NOTE: Linguists sometimes say that “construction X expresses construction Y”, e.g.

“the relative clause is expressed by a participial construction”

“a juxtaposition construction expresses the coordination construction”

“the possessive construction is expressed by a person affix on the noun”

Here, X is a construction-strategy, and Y is a construction-function.

These ways of talking are not reversible (*the participial construction is expressed by a relative clause construction) – strategies express construction-functions, but not the other way round.

In the past, the distinction between construction-functions and construction-strategies was often formulated in terms of “syntax” and “morphology”, e.g. in Lander & Nichols (2020):

“Head/dependent marking typologizes word and phrase morphosyntax, constructions, and whole languages according to whether grammatical relations are overtly marked on the head of the relevant phrase, a non-head, both, or neither. ...

(i) a. Hungarian
az ember ház-a
 the man house-3SG.POSS
 ‘the man’s house’

b. English
the man’s house

The essential theoretical point is that the **syntax** of the Hungarian and English phrases is identical (at least to the extent that it is comparable) but the **morphology** is opposite.”

Here it would be better to say that the **adnominal possessive construction** (functionally defined) can be expressed by an **indexing strategy** (as in Hungarian) or by a **flagging strategy** (as in English).

(Indexing and flagging are construction types, too – and the “syntax” vs. “morphology” distinction is hardly relevant.)

The terminology used by linguists is often confusing – one and the same term is often used for both **functions** and **strategies**.

(1) **reflexive construction** (cf. Haspelmath 2023a)

stereotype: *They saw **themselves** in the mirror.*
 • form-based: *They behaved **themselves**.*
 • function-based: *They dressed.* (cf. Reuland 2011: “reflexive predicate”)

(2) **adjectival construction**

stereotype: (English) *the **new** house*
 • form-based: (English) *the royal family*
 • function-based: (Mandarin) *fángzi hěn dà* ‘the house is big’ (predicative Verb)

(3) **nominalization construction**

stereotype: *the construct-**ion** of the city*
 • form-based: *the construct-ions in French*
 • function-based: [*that the city was constructed*] (cf. Shibatani 2019)

(4) **impersonal construction** (cf. Malchukov & Ogawa 2011)

stereotype: (Latin) *curritur* ‘there is running’ (impersonal passive)
 • form-based: (Spanish) *llueve* ‘it is raining’, no subject
 • function-based: (French) *on court* ‘one runs’, no specified agent

For a construction-function, a typology may ask about **types of strategies**:

- e.g. types of *polar questions* (Dryer 2005c)
 types of *predpossessive clauses* (Stassen 2005)
 types of *ordinal numerals* (Stolz & Veselinova 2005)
 types of *relative clauses* (Dryer 2005b)
 types of *plural number* constructions (Dryer 2005a)

For a construction-strategy, a typology may ask about additional meanings (**coexpression patterns**):

- e.g. add. meanings of *reciprocal markers* (e.g. reflexive; Maslova & Nadjalkov 2005)
 add. meanings of *comitative flags* (e.g. instrumental; Stolz et al. 2005)
 add. meanings of *relativizers* (e.g. genitive flag; Gil 2005)
 add. meanings of *nominal conjunctors* (e.g. verbal conj.; Haspelmath 2005b)

II. Efficiency explanations

II.1. Types of explanations for universals

We saw that if we want to engage in **general linguistics**, we must study **universals**. The study of a particular language (p-linguistics) gives us insight into this language but not (necessarily) about Human Language.

Once we have found language universals, we can explain them in different ways (Haspelmath 2019):

- **functional-adaptive:** languages are (culturally) adapted to their users' needs
 (e.g. efficient coding, domain minimization)
- **biocognitive-representational:** languages are limited by what can be (biocognitively) represented in speakers's minds
 (e.g. "grammars don't count")
- **mutational:** certain changes are impossible or unlikely
 (e.g. "direction markers don't become comitative markers")

The best approach is to consider all the possibilities, and choose the most likely explanation.

II.2. Universals of asymmetric coding: Simple meaning pairs

Haspelmath (2021)

Languages often have pairs of grammatical meanings, one of which is expressed by an overt (or a longer) form. These **coding asymmetries** are often cross-linguistically general.

Table 1: Examples of universal grammatical coding asymmetries

present	future	(<i>go</i> – <i>will go</i>)
singular	plural	(<i>book</i> – <i>book-s</i>)
3 rd person	2 nd person	(Spanish <i>canta</i> – <i>canta-s</i>)
nominative	accusative	(Hungarian <i>ember</i> – <i>ember-t</i>)
affirmative	negative	(<i>go</i> – <i>don't go</i>)
allative	ablative	(<i>to</i> – <i>from</i>)
positive	comparative	(<i>small</i> – <i>small-er</i>)

What do the meanings in the right column have in common?

Traditionally: they are **semantically more complex**.

A better explanation: they are **more frequent** (Haspelmath 2008a)

2.1. Singular vs. plural

	Hebrew	Khanty
SG	<i>yom</i>	<i>xot</i>
PL	<i>yam-im</i>	<i>xot-at</i>
	‘day(s)’	‘house(s)’

2.2. Nominative vs. accusative

	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>

Each of these examples represents a universal trend, e.g.

In the world’s languages, if the coding is asymmetric, the singular marker is generally shorter than the plural marker (and often zero).

In the world’s languages, if the coding is asymmetric, the nominative marker is generally shorter than the accusative marker.

2.3. Allative vs. ablative

	English	Sri Lanka	Portuguese	Japanese
ALLATIVE	<i>to Rome</i>	<i>maaket</i> ‘to the market’		<i>Tookyoo e</i> ‘to Tokyo’
ABLATIVE	<i>from Rome</i>	<i>kaaza impa</i> ‘from home’		<i>Tookyoo kara</i> ‘from Tokyo’

2.4. Male vs. female occupational terms

	Latin	German	Hungarian	
MALE	<i>rex</i>	<i>König</i>	<i>király</i>	‘king’
FEMALE	<i>reg-ina</i>	<i>König-in</i>	<i>király-nő</i>	‘queen’

2.5. Present tense vs. future tense

	English	Latin	Kiribati
PRS	<i>they praise</i>	<i>lauda-nt</i>	<i>e taetae</i> ‘he speaks’
FUT	<i>they will praise</i>	<i>lauda-b-unt</i>	<i>e na taetae</i> ‘he will speak’

In all these cases, the longer forms are less frequent, in all languages. In general, we can say:

(5) The grammatical form-frequency correspondence hypothesis

When two grammatical construction types that differ minimally (i.e. that form a semantic opposition) occur with significantly different frequencies, the less frequent construction tends to be overtly coded (or coded with more segments), while the more frequent construction tends to be zero-coded (or coded with fewer segments), if the coding is asymmetric.

The idea of form-frequency correspondences in grammar can be traced back to Greenberg (1966) (see also Croft 2003: Ch. 4; Hawkins 2004: §3.2.2; Haspelmath 2008a; 2008b; Diessel 2019: Ch. 11).

(2) The form-frequency correspondence universal

Languages tend to have shorter forms for more frequent meanings.

This is an old insight and is uncontroversial for word length (e.g. Zipf 1935: 23).

The proposed explanation is **coding efficiency**: More frequently expressed meanings are **more predictable** and therefore need less coding. **This is efficient.**

causal chain:

frequency of use → predictability → coding by zero or a short form

Speakers tend to expend only as much articulatory energy as necessary, so when a construction type is frequent, they can economize. Over time, speakers’ discourse tendencies may get **grammaticalized** in languages.

II.3. Universals of asymmetric coding: 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) Spanish

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) Hebrew

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.

There is also 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-sin</i>	‘let her look’
INDICATIVE	2ND	<i>lauda-v-isti</i>	‘you praised’	<i>bak-tyor-un</i>	‘you are looking’
	3RD	<i>lauda-v-it</i>	‘she praised’	<i>bak-tyor-Ø</i>	‘she is looking’

There is also a **usual association** of semantic class with discourse function:

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

As a result, in most languages, property concept roots have short (or Ø) coding in **attributive function**, while 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).

		French	M. Chinese	
ATTR	PROPERTY	<i>petit Ø enfant</i>	<i>xiǎo Ø háizi</i>	‘small child’
	ACTION	<i>enfant qui joue</i>	<i>wán de háizi</i>	‘child who plays’
PRED	PROPERTY	<i>l’enfant est petit</i>	<i>háizi Ø xiǎo</i>	‘the child is small’
	ACTION	<i>l’enfant joue-Ø</i>	<i>háizi Ø wán</i>	‘the child plays’

II.4. Conclusion

Like universals of word order (Hawkins 2014), universals of asymmetric coding can be explained as a **functional-adaptive response** to demands on speaker and hearers. It is much more efficient to use overt coding when it is unexpected than the other way round.

Maybe the reason why processing explanations of word order have been more widely accepted and discussed is that

- (i) it is intuitively clear that some word orders are difficult to process
- (ii) meanings do not play a big role in ordering words

But **efficiency explanations** have been successful in a wide range of domains (e.g. Gibson et al. 2019), and nobody really doubts the efficiency explanation for Zipfian shortness of words (“Short words are short because they are frequent”).

If we stop thinking only in terms of boxes (grammar vs. lexicon etc.), it is very natural to extend the efficiency explanation also to grammatical asymmetries of coding.

References

- Aikhenvald, Alexandra Y. 2010. *Imperatives and commands*. Oxford: Oxford University Press.
- Croft, William. 1991. *Syntactic categories and grammatical relations: The cognitive organization of information*. Chicago: University of Chicago Press.
- Croft, William. 2003. *Typology and universals*. 2nd edition. Cambridge: Cambridge University Press.
- Croft, William. 2022. *Morphosyntax: Constructions of the world’s languages*. Cambridge: Cambridge University Press.
- Diessel, Holger. 2019. *The grammar network*. Cambridge: Cambridge University Press.
- Geniušienė, Emma. 1987. *The typology of reflexives*. Berlin: Mouton de Gruyter.
- Gibson, Edward & Futrell, Richard & Piandadosi, Steven T. & Dautriche, Isabelle & Mahowald, Kyle & Bergen, Leon & Levy, Roger. 2019. How efficiency shapes human language. *Trends in Cognitive Sciences* 23(5). 389–407. (doi:[10.1016/j.tics.2019.02.003](https://doi.org/10.1016/j.tics.2019.02.003))
- Haiman, John. 1983. Iconic and economic motivation. *Language* 59(4). 781–819. (doi:[10.2307/413373](https://doi.org/10.2307/413373))
- Haspelmath, Martin. 2008a. 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. 2008b. Creating economical morphosyntactic patterns in language change. In Good, Jeff (ed.), *Linguistic universals and language change*, 185–214. Oxford: Oxford University Press.
- Haspelmath, Martin. 2019. Can cross-linguistic regularities be explained by constraints on change? In Schmidtke-Bode, Karsten & Levshina, Natalia & Michaelis, Susanne Maria & Seržant, Ilja A. (eds.), *Competing explanations in linguistic typology*, 1–23. Berlin: Language Science Press. (<http://langsci-press.org/catalog/book/220>)
- Haspelmath, Martin. 2021. Explaining grammatical coding asymmetries: Form-frequency correspondences and predictability. *Journal of Linguistics* 57(3). 605–633. (doi:[10.1017/S0022226720000535](https://doi.org/10.1017/S0022226720000535))
- Haspelmath, Martin. 2023a. Comparing reflexive constructions in the world’s languages. In Janic, Katarzyna & Puddu, Nicoletta & Haspelmath, Martin (eds.), *Reflexive constructions in the world’s languages*, 19–62. Berlin: Language Science Press.
- Haspelmath, Martin. 2023b. On what a construction is. *Constructions* 15(1). (doi:[10.24338/cons-539](https://doi.org/10.24338/cons-539))
- Hawkins, John A. 2004. *Efficiency and complexity in grammars*. Oxford: Oxford University Press.

- Hawkins, John A. 2014. *Cross-linguistic variation and efficiency*. New York: Oxford University Press.
- Malchukov, Andrej L. & Ogawa, Akio. 2011. Towards a typology of impersonal constructions: A semantic map approach. In Malchukov, Andrej & Siewierska, Anna (eds.), *Impersonal constructions: A cross-linguistic perspective*, 19–56. Amsterdam: Benjamins.
- Shibatani, Masayoshi. 2019. What is nominalization? Towards the theoretical foundations of nominalization. In Zariquiey, Roberto & Shibatani, Masayoshi & Fleck, David W. (eds.), *Nominalization in languages of the Americas*, 15–167. Amsterdam: Benjamins.
- Siewierska, Anna. 2010. Person asymmetries in zero expression and grammatical function. In Floricic, Franck (ed.), *Essais de typologie et de linguistique générale: Mélanges offerts à Denis Creissels*, 471–485. Lyon: ENS Éditions.
- Zipf, George Kingsley. 1935. *The psycho-biology of language: An introduction to dynamic philology*. Cambridge: MIT Press.

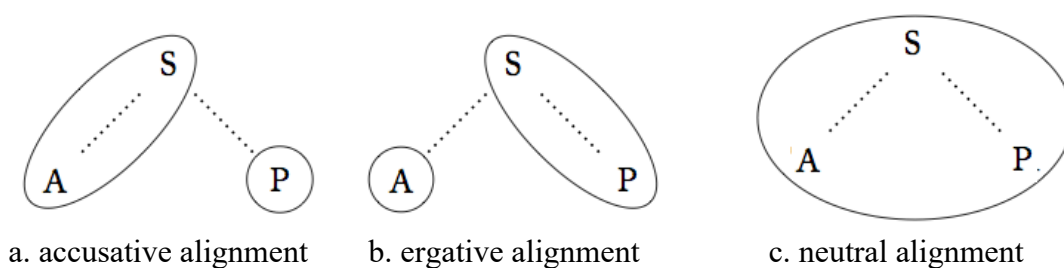
Alignment and differential coding:

2A. Accusative and ergative alignment

MARTIN HASPELMATH

Max Planck Institute for Evolutionary Anthropology

1. Accusative, ergative and neutral alignment (in monotransitive constructions)

**Figure 1: The three monotransitive alignment types**

accusative pattern: transitive P (patient) gets a special marker (*accusative* marker)
transitive A (agent) is marked like the intransitive subject (S)

(1) Latin (Napoli 2018: 64, 62)

a. *ille fugit periculo*
he.NOM escaped danger
'he escaped from danger'

b. *ill-um ex periculo eripuit*
he-ACC from danger rescued.3SG
'he rescued him from danger' (Caes. *Gall.* 4, 12, 5)

ergative pattern: transitive A (agent) gets a special marker (*ergative* marker)
transitive P (patient) is marked like the intransitive subject (S)

(2) Kavalan (a language of Taiwan; Liao 2004: 214-215)

a. *Mawtu=ti a paqapaRan ta taiwan.*
come=PRF ABS catcher LOC Taiwan
'The catcher came to Taiwan.'

b. *Inebana=ti na hetay a rawang.*
close=PRF ERG soldier ABS city
'The soldiers closed the city (gate).'

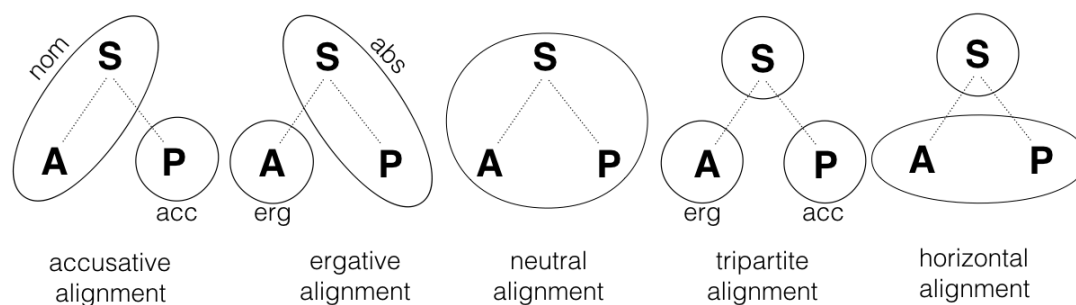


Figure 2. Five alignment types of argument markers

Universal 1: The accusative-ergative universal

In all languages, if the coding is asymmetric, accusative and ergative flags are longer than nominative and absolutive flags (Comrie 1978; Dixon 1979).

Explanation: frequency-induced predictability

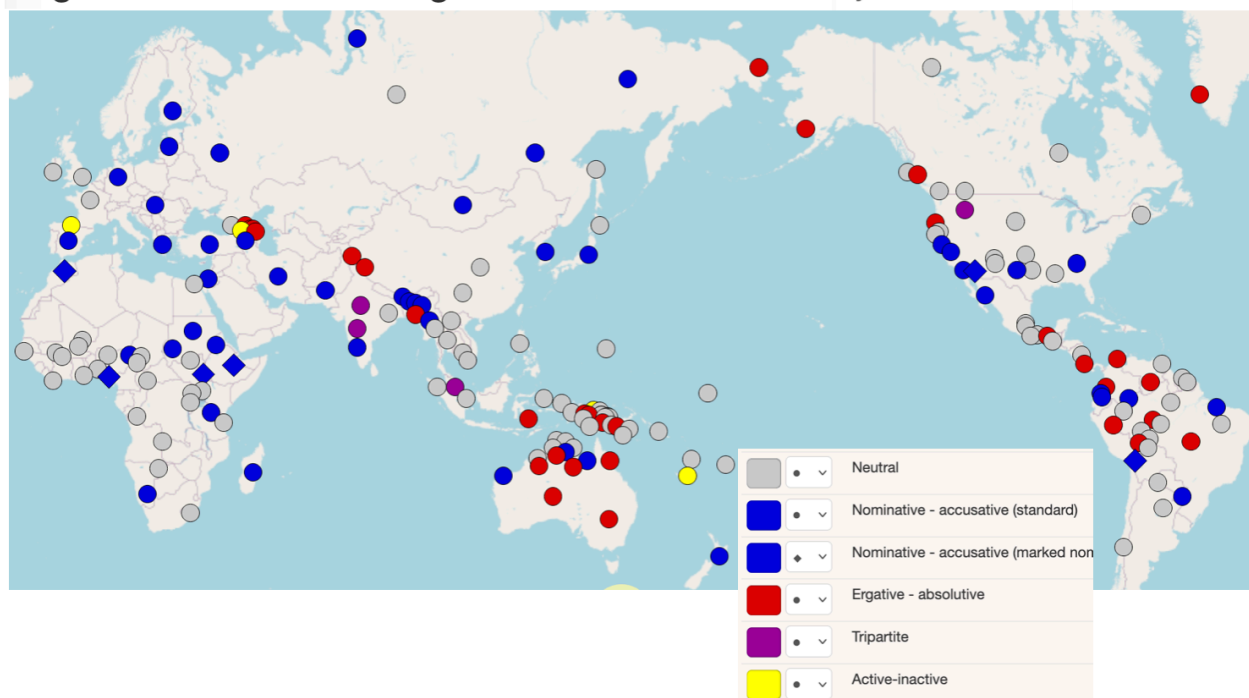
- because accusative is less frequent than nominative, and ergative is less frequent than absolutive

Moreover, the only commonly occurring patterns are accusative, ergative and neutral.

This has a well-known functional explanation:

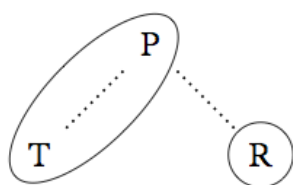
- tripartite alignment is **uneconomical**, and horizontal alignment **does not make the right distinctions**

Alignment of Case Marking of Full Noun Phrases by Bernard Comrie

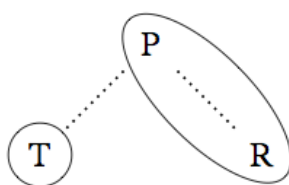


2. Indirective, secundative and neutral alignment (in ditransitive constructions)

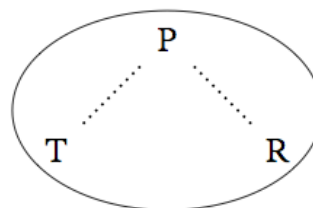
Similarly, there are three types of **ditransitive alignment** (Malchukov et al. 2010; Haspelmath 2015):



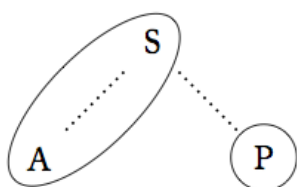
a. indirective alignment
(P=T vs. R)



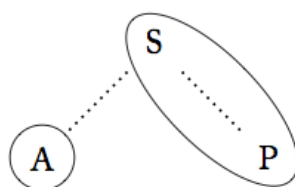
b. secundative alignment
(P=R vs. T)



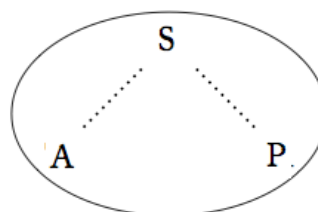
c. neutral alignment
(P=T=R)



a. accusative alignment



b. ergative alignment



c. neutral alignment

(2) indirective alignment: French

a. *Marie mange la soupe.*

Marie eats the soup

P

‘Marie is eating the soup.’

b. *Marie donne la soupe à sa mère.*

Marie gives the soup to her mother

T

R

‘Marie is giving the soup to her mother.’

(3) secundative alignment: Yoruba (Rowlands 1969: 20-21)

a. *Ó pè mí.*

she call me

P

‘She called me.’

b. *Ó fún mí l’ ówó.*

he give me SEC money

R

T

‘He gave me money.’

(4) neutral alignment: Diyari (Austin 1981: 115)

- a. *nhulu wanku-yali mankarra-wurla-nha nyayinyayi-yi*
 ART.NF.ERG snake-ERG girl-DL-ACC watch-PRS
 P
 'The snake watched the girls.'
- b. *nhulu pulanha nhinha putu yingki-rna wara-yi*
 3SG.NF.ERG 3DL.ACC ART.NF.ACC thing.ACC give-PTCP AUX-PRS
 R T
 'He gave them (two) that thing.'

Universal 2: The indirective-secundative universal

In all languages, if the coding is asymmetric, indirective and secundative markers are longer than directive and primitive markers (Malchukov et al. 2010; Haspelmath 2015).

Two more examples:

(5) a. R flag is overt, no T flag

Wappo (California)

ce k'ew-i chica-thu ew ma-hes-ta?
 DEM man-NOM bear-DAT fish DIR-give-PST
 'The man gave the fish to the bear.' (Thompson et al. 2006: 12)

b. T flag is overt, no R flag

Teiwa (Alor-Pantar)

Ui ga'an u sen ma n-oma' g-an.
 person 3SG DIST money SEC 1SG.POSS-father 3SG.OBJ-give
 'That person gave money to my father.' (Klamer 2010: 444)

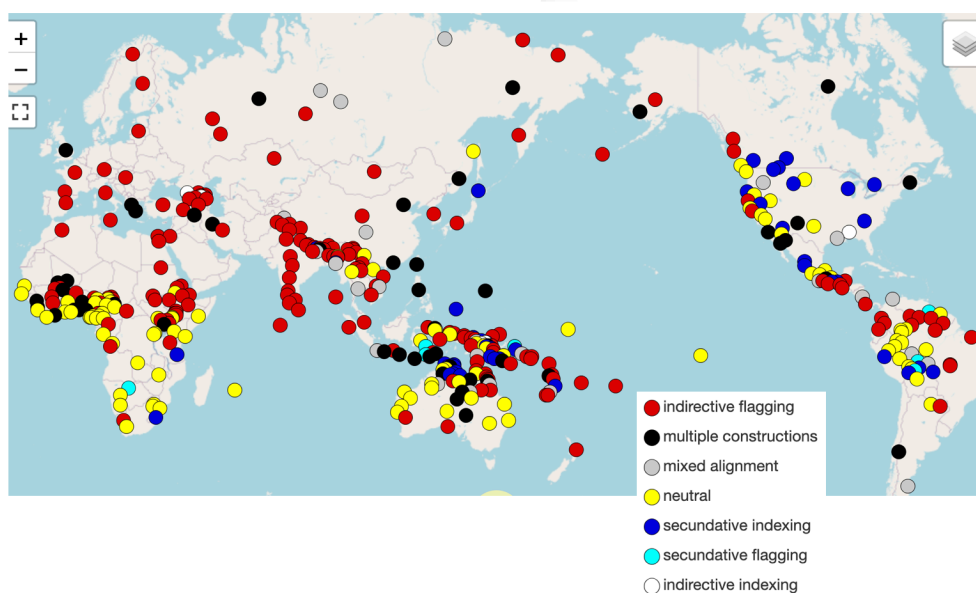
Explanation: as above (frequency-induced predictability)

Three striking features of these universals:

- they are based on the notion of **flag** (a **concrete form**), rather than on abstract “case assignment”
- they are based on a view of core argument marking in which S=A alignment (accusative) and S=P alignment (ergative) **are treated on a par** – accusative and ergative are mirror-image concepts, with no preference for the European-style accusative pattern
- they also highlight the parallelism between **monotransitive** and **ditransitive** constructions:

The asymmetry of **A and P** in monotransitives is parallel to the asymmetry of **R and T**!

dominant ditransitive construction by Andrej Malchukov and Martin Haspelmath



3. Alignment of indexing

Accusative vs. ergative alignment can also be shown by indexing, and this also applies to indirective vs. secundative alignment.

(6) English (**accusative indexing: S=A**)

- The tree fall-s.*
- The mother hug-s the boy.*

(-s : “nominative index”)

(S indexed)

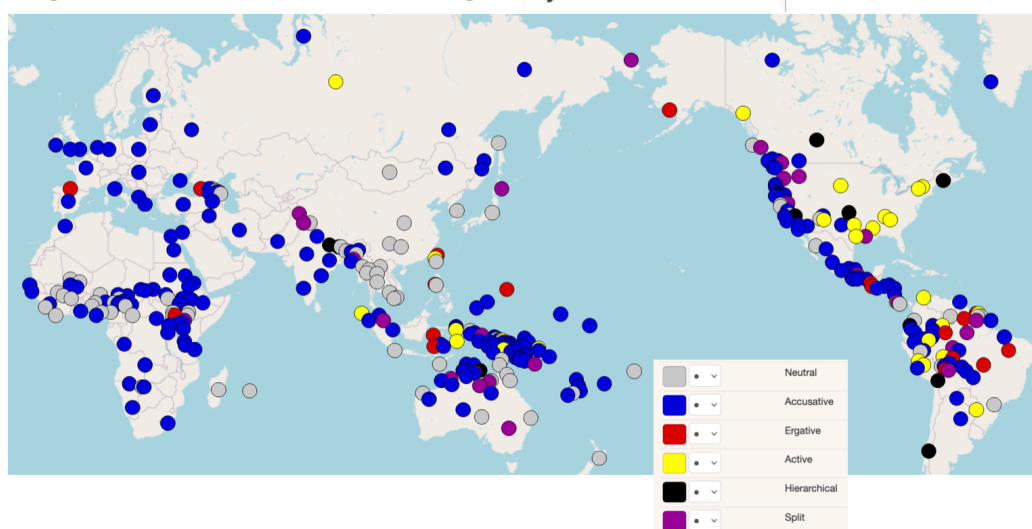
(A indexed)

(7) Yucatec Maya (Lehmann 2015; **ergative indexing: S=P**)

- H=liub-Ø le che'-o'.* (S indexed)
 PFV=fall.CMPL-**3SG.ABS** DEM wood-DEM2
 ‘The tree fell.’

- T-u=méek'-ah-Ø u=chaan xibpal le maamah-o'.*
 PFV-**3.ERG**=hug-CMPL-**3SG.ABS** 3.POSS=little boy DEM mother-DEM2
 ‘The mother hugged her little boy.’ (P indexed)

Alignment of Verbal Person Marking by Anna Siewierska



Whatever the explanation, it is clear that the **monotransitive A** and the **ditransitive R** are similar:

A and R tend to be human and definite, while P and T tend to be nonhuman and indefinite – we can say that A and R are “higher in role rank”:

A > P

R > T

4. Are ergative clauses transitive? Comments on Mel'čuk

This depends on the definition of “transitivity”...

Igor A. Mel'čuk (*1932) was a visiting professor at the University of Vienna in 1983, when I was a first-year student there.

There I learned about Mel'čuk (1981), a paper about Lezgian, with a fascinating theory of Lezgian clause structure:

Mel'čuk claimed that Lezgian has no transitive ergative construction, but that its Ergative-marked argument, as seen in (12), is an oblique causal complement!

- (12) Алиди кицІ кьена.
Ali-di kic' q'e-na.
 Ali-ERG dog kill-AOR
 ‘Ali killed the dog.’

Basically, **all Lezgian clauses are claimed to be intransitive**. Clause (1) literally means ‘The dog died through Ali’. (See also Mel'čuk 1988; 2013.)

In one of my first journal papers (Haspelmath 1991), I tried to show that this was wrong: (12) was in fact a transitive clause – but how could I tell?

What is a “transitive clause”, in general?

According to Mel'čuk an *ergative construction* must have ergative marking on the *subject* – and he is not convinced that the ergative nominal in (12) is a subject.

But what is a “subject”, in general?

Many authors define “subject” in such a way that **different criteria** are applied in **different languages**:

Mel’čuk (2013: x):

“...the SyntSubj [is] cross-linguistically universal. However, in a different sense, the SyntSubj is **language-specific in so far as syntactic privileges are different in different languages**: thus, in many Indo-European languages the main privilege of a clausal element is to impose agreement on the Main Verb, while in Malagasy it is to occupy the clause-final position.”

Riesberg et al. (2019: 524)

“We use the term ‘subject’ here as equivalent to what is termed ‘privileged syntactic argument’ (PSA) in Van Valin (2005) and elsewhere. A PSA is defined as the syntactic element that **controls coding properties *such as* agreement and that is the pivotal element in complex constructions *such as* relativization, NP deletion, control, and so forth.**”

But if different subjects may be recognized by different criteria, how do we know that they are all “subjects”? (By persuasion? cf. Haspelmath 1991)

Another example of an unexpected view of transitivity: Legate (2012: 184):

“Warlpiri ... allows two-argument verbs with an ergative-dative case frame, indicating that a dative DP can satisfy the transitivity restriction.”

- (5) a. ERG-DAT
 Ngarrka-ngku-rla-jinta marlu-ku pantu-rnu, [marna nga-rninja-kurra(-ku)].
 man-ERG-3DAT-DAT kangaroo-DAT spear-PAST, grass-ABS eat-*INFIN-C*-(DAT)
 ‘The man speared at the kangaroo eating grass.’ (Simpson and Bresnan, 1983:54)

Legate refers to “a well-established test for objecthood in Warlpiri” (the use of particular subordinators), but such tests are not applicable to languages in general, so they do not lead to general conclusions.

5. The argument-types S, A and P: Comparative concepts for argument classes (“syntactic functions”)

Comparison of languages must be based on uniform “units of measurement” – in other words, the concepts which we use to compare languages must be identified **in the same way in all languages** (Haspelmath 2010; 2018a)

Comrie (1978), Lazard (2002), Haspelmath (2011):

- A is the agent argument of a physical-effect verb like ‘kill’ or ‘break’ (in the usual construction), plus verbs with the same coding
- P is the patient argument of such verbs
- S is the argument of a change-of-state verb like ‘fall’ or ‘die’, plus verbs with the same coding

These comparative concepts can be applied to all languages **uniformly** (= using the **same criteria**), because all languages have verbs with these meanings and have nominal arguments with particular coding properties.

(Note: S, A and P are not “macroroles”, in the sense of Dowty 1991, or Role and Reference Grammar!)

Alternatively, one might also use more **fine-grained semantic roles** (verb-specific roles, “microroles”), as in Hartmann et al. (2014) (based on ValPaL):

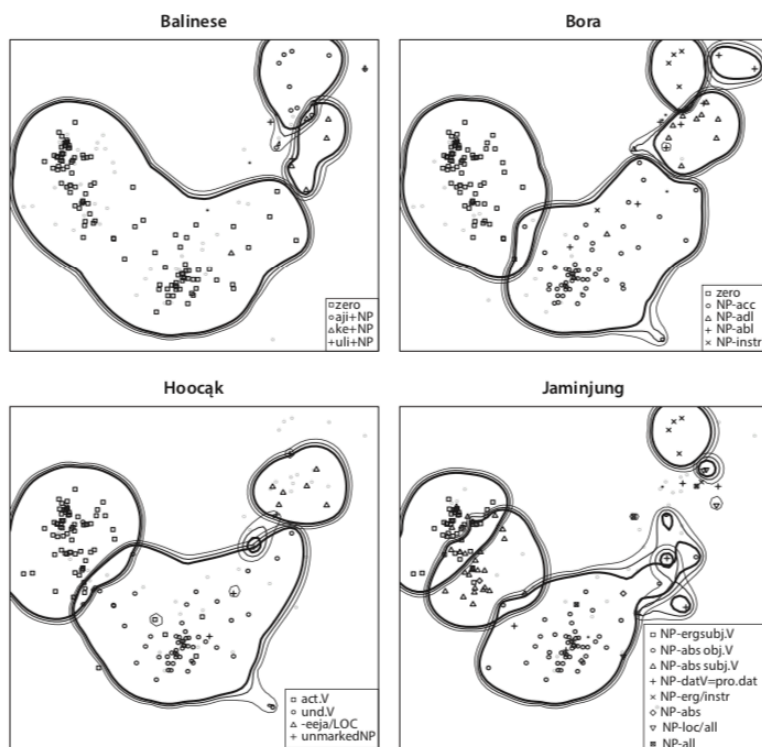


Figure 5. Four additional languages showing different distributional ranges of coding sets

These images show that verbs cluster in all languages, and they do so in a way that is cross-linguistically similar. The use of S, A and P is a simplification, but it is useful for many purposes.

References

- Aissen, Judith. 2003. Differential object marking: Iconicity vs. economy. *Natural Language & Linguistic Theory* 21(3). 435–483.
- Austin, Peter K. 1981. *A grammar of Diyari, South Australia*. Cambridge: Cambridge University Press.
- Baker, Mark C. & Bobaljik, Jonathan. 2017. On inherent and dependent theories of ergative case. In Coon, Jessica & Massam, Diane & Travis, Lisa deMena (eds.), *The Oxford handbook of ergativity*. Oxford: Oxford University Press.
- Baker, Mark C. 2001. *The atoms of language*. New York: Basic Books.
- Baker, Mark C. 2015. *Case: Its principles and parameters*. Cambridge: Cambridge University Press.
- Blake, Barry J. 1976. On ergativity and the notion of subject: Some Australian cases. *Lingua* 39(4). 281–300. (doi:10.1016/0024-3841(76)90048-6)

- Comrie, Bernard. 1978. Ergativity. In Winfred P. Lehmann (ed.), *Syntactic typology: Studies in the phenomenology of language*, 329–394. Austin: University of Texas Press.
- Croft, William. 2022. *Morphosyntax: Constructions of the world's languages*. Cambridge: Cambridge University Press. (<http://www.unm.edu/~wcroft/WACpubs.html>)
- Deal, Amy Rose. 2015. Ergativity. In Kiss, Tibor & Alexiadou, Artemis (eds.), *Syntax: Theory and analysis (Volume 1)*, 654–708. Berlin: De Gruyter Mouton. (doi:10.1515/9783110377408.654)
- DeLancey, Scott. 1981. An interpretation of split ergativity and related patterns. *Language* 57. 626–657.
- Dixon, R.M.W. 1979. Ergativity. *Language* 55. 59–138.
- Foley, William A. & Van Valin, Robert D. 1977. On the viability of the notion of “subject” in universal grammar. *Berkeley Linguistics Society* 3. 293–320.
- Hartmann, Iren, Martin Haspelmath & Michael Cysouw. 2014. Identifying semantic role clusters and alignment types via microrole coexpression tendencies. *Studies in Language* 38(3). 463–484.
- Haspelmath, Martin. 1991. On the question of deep ergativity: The evidence from Lezgian. 44/45(1–2). 5–27. doi:10.5281/zenodo.225289.
- Haspelmath, Martin. 1993. *A grammar of Lezgian* (Mouton Grammar Library 9). Berlin: Mouton de Gruyter.
- Haspelmath, Martin. 2007. Ditransitive alignment splits and inverse alignment. *Functions of Language* 14(1). 79–102.
- Haspelmath, Martin. 2010. Comparative concepts and descriptive categories in crosslinguistic studies. *Language* 86(3). 663–687. doi:10.1353/lan.2010.0021.
- Haspelmath, Martin. 2011. On S, A, P, T, and R as comparative concepts for alignment typology. *Linguistic Typology* 15(3). 535–567.
- Haspelmath, Martin. 2015. Ditransitive constructions. *Annual Review of Linguistics* 1. 19–41. (doi:10.1146/annurev-linguist-030514-125204)
- Haspelmath, Martin. 2018a. How comparative concepts and descriptive linguistic categories are different. In Daniël Van Olmen, Tanja Mortelmans & Frank Brisard (eds.), *Aspects of linguistic variation: Studies in honor of Johan van der Auwera*, 83–113. Berlin: De Gruyter Mouton. <http://doi.org/10.5281/zenodo.570000>.
- Haspelmath, Martin. 2018b. Review of “Baker, Mark. 2015. *Case*. Cambridge: Cambridge University Press.” *Studies in Language* 42(2). 474–486. (doi:<https://doi.org/10.1075/sl.16059.has>)
- Haspelmath, Martin. 2019a. Indexing and flagging, and head and dependent marking. *Te Reo* 62(1). 93–115. doi:10.17617/2.3168042.
- Haspelmath, Martin. 2019b. Ergativity and depth of analysis. *Rhema* 2019(4). 108–130. (doi:10.31862/2500-2953-2019-4-108-130)
- Haspelmath, Martin. 2021. Role-reference associations and the explanation of argument coding splits. *Linguistics* 59(1). 123–174. (doi:10.1515/ling-2020-0252)
- Jackendoff, Ray. 2007. *Language, consciousness, culture: Essays on mental structure* (Jean Nicod Lectures). Cambridge, Mass.: MIT Press. (<http://site.ebrary.com/id/10190461>) (Accessed June 25, 2023.)
- Kibrik, Alexandr (ed.). 1996. *Godoberi*. München: Lincom Europa.
- Lazard, Gilbert. 2002. Transitivity revisited as an example of a more strict approach in typological research. *Folia Linguistica* 36(3–4). 141–190.
- Legate, Julie Anne. 2012. Types of ergativity. *Lingua* (Accounting for Ergativity) 122(3). 181–191. (doi:10.1016/j.lingua.2011.10.014)
- Liao, Hsiu-chuan. 2004. *Transitivity and ergativity in Formosan and Philippine languages*. Manoa: University of Hawai’i at Manoa. (PhD dissertation.)
- Malchukov, Andrej & Haspelmath, Martin & Comrie, Bernard. 2010. Ditransitive constructions: A typological overview. In Malchukov, Andrej & Haspelmath, Martin & Comrie, Bernard (eds.), *Studies in ditransitive constructions*, 1–64. Berlin: Mouton de Gruyter.

- Mel'čuk, Igor A. 2013. Syntactic subject, once again. In Valentina Apresjan, Boris Iomdin & Ekaterina Ageeva (eds.), *Proceedings of the 6th International Conference on Meaning-Text Theory, Prague, August 30-31, 2013*, iii–xxxiii. Prague.
- Mel'čuk, Igor A. 1981. Grammatical subject and the problem of the ergative construction in Lezgian. In Bernard Comrie (ed.), *Studies in the Languages of the USSR*, 229–276. Carbondale & Edmonton: Linguistic Research.
- Mel'čuk, Igor A. 1988. *Dependency syntax: Theory and practice*. Albany: State University Press of New York.
- Moravcsik, Edith A. 1978. On the distribution of ergative and accusative patterns. *Lingua* 45. 233–279.
- Napoli, Maria. 2018. Ditransitive verbs in Latin: A typological approach. *Journal of Latin Linguistics* 17(1). 51–91. (doi:10.1515/joll-2018-0003)
- Nichols, Johanna. 2011. *Ingush grammar* (University of California Publications in Linguistics, 141). Berkeley: University of California Press.
- Petroski, Henry. 1994. *The evolution of useful things*. Vintage. New York: Vintage Books. (<http://catdir.loc.gov/catdir/enhancements/fy1102/93006351-b.html>) (Accessed June 25, 2023.)
- Reiss, Charles. 2018. Substance-free phonology. In Hannahs, Stephen J. & Bosch, Anna R. K. (eds.), *The Routledge handbook of phonological theory*. London: Routledge.
- Riesberg, Sonja, Kurt Malcher & Nikolaus P. Himmelmann. 2019. How universal is agent-first? Evidence from symmetrical voice languages. *Language* 95(3). 523–561. doi:10.1353/lan.2019.0055.
- Rowlands, E. C. 1969. *Teach yourself Yoruba*. London: Hodder and Stoughton.
- Schmidtke-Bode, Karsten & Natalia Levshina. 2018. Assessing scale effects on differential case marking: Methodological, conceptual and theoretical issues in the quest for a universal. In Ilja A. Seržant & Alena Witzlack-Makarevich (eds.), *Diachrony of differential argument marking* (Studies in Diversity Linguistics 19), 463–489. Berlin: Language Science Press.
- Silverstein, Michael. 1976. Hierarchy of features and ergativity. In Dixon, Robert M. W. (ed.), *Grammatical categories in Australian languages*, 112–171. Canberra: Australian National University.
- Simpson, Jane & Bresnan, Joan. 1983. Control and obviation in Warlpiri. *Natural Language and Linguistic Theory* 1(1). 49–64.
- Skirgård, Hedvig & Haynie, Hannah J. & Blasi, Damián E. & Hammarström, Harald & Collins, Jeremy & Latache, Jay J. & Lesage, Jakob et al. 2023. Grambank reveals the importance of genealogical constraints on linguistic diversity and highlights the impact of language loss. *Science Advances* 9(16). eadg6175. (doi:10.1126/sciadv.adg6175)
- Van Valin, Robert D., Jr. 2005. *Exploring the syntax-semantics interface*. Cambridge: Cambridge University Press.
- Vaux, Bert & Myler, Neil. 2018. Outstanding issues and future prospects in Rule-Based Phonology. In Hannahs, Stephen J. & Bosch, Anna R. K. (eds.), *The Routledge handbook of phonological theory*. London: Routledge.

2. The Lezgian language (and my professor I.A. Mel'čuk)

Igor A. Mel'čuk (*1932) was a visiting professor at the University of Vienna in 1983, when I was a first-year student there.

There I learned about Mel'čuk (1981), a paper about Lezgian, with a fascinating theory of Lezgian clause structure:

Mel'čuk claimed that Lezgian has no transitive ergative construction, but that its Ergative-marked argument, as seen in (1), is an oblique causal complement!

- (1) Алиди кицІ кьена
Ali-di kic' q'e-na.
 Ali-ERG dog kill-AOR
 'Ali killed the dog.'

Basically, all Lezgian clauses are claimed to be intransitive. Clause (1) literally means 'The dog died through Ali'. (See also Mel'čuk 1988; 2013.)

In one of my first journal papers (Haspelmath 1991), I tried to show that this was wrong. (1) was in fact a transitive clause – but how could I tell?

What is a “transitive clause”, in general?

Mel'čuk provides detailed discussions of the terms *subject* and *ergative construction*, and I learned a lot from his careful approach to what our grammatical terms mean.

3. Differential ergative and accusative marking

Differential marking is often conditioned by **referential prominence**:

differential A-marking is found primarily with low-prominence A-arguments

- 3rd person (vs. locuphoric, i.e. 1st/2nd person)
- inanimate (vs. animate)
- indefinite/focus (vs. definite/topic)

- (2) Godoberi (Nakh-Dagestanian; Kibrik (ed.) 1996: 108))

a. *imu-di ʕali č'inni*
 father-ERG Ali(ABS)beat.PST
 'Father beat Ali.'

b. *min ʕali č'inni*
 you Ali(ABS) beat.PST
 'You beat Ali.'

c. *imu-di min č'inni*
 father-ERG you beat.PST
 'Father beat you.'

This pattern is a universal tendency (cf. DeLancey 1981; Testelec 1986; Schmidtke-Bode & Levshina 2018), and it can be generalized as follows:

- (3) **The role-reference association universal** (Haspelmath 2021b)

Deviations from usual associations of role rank and referential prominence tend to be coded by longer grammatical forms if the coding is differential.

e.g. differential P marking for definite or animate nominals

e.g. differential (ditransitive) R marking for indefinite nominals (Haspelmath 2007)
etc.

A and R are usually associated with **high** referential prominence.

P and T are usually associated with **low** referential prominence.

This is a frequency effect:

“**Usual association**” means that these associations are the most frequent ones, and hence the most **predictable** – which means that it is efficient if they get **less coding** than the less predictable meanings.

Haspelmath (2021a): Efficiency theory of asymmetric coding in grammar.

Thus, quite a few initially puzzling aspects of the grammar of ergativity have a good functional-adaptive explanation.

3. “Depth of analysis” and measurement uniformity

e.g. Polinsky & Kluender (2007)

“[Typology’s] allegiance to large samples and “superficial” generalizations is simply one of the consequences of casting the net wide and looking for differences in a quick and easy way.”

e.g. Polinsky (2011)

“Such research allows us to uncover subtle distinctions and fine details of grammar that often remain unnoticed in a coarse-grained approach to language typology.”

e.g. Baker (2015: 287)

“This testifies to the deep unity of human language, discernable underneath the surface diversity of case patterns.”

e.g. Roberts (2019: 12):

“From the perspective of generative grammar, much typological analysis seems excessively surface-oriented.”

I would like to argue that this very widespread view is mistaken. The relative negligence of “depth of analysis” in comparative research is not a bug – it’s a necessary feature.

Comparison of languages must be based on uniform “units of measurement” – in other words, the concepts which we use to compare languages must be identified **in the same way in all languages**.

This is what I call “**measurement uniformity**” (and I will briefly contrast it with “building block uniformity” later).

For example, how are S, A and P defined?

Comrie (1978), Lazard (2002), Haspelmath (2011a):

- A is the agent argument of a physical-effect verb like ‘kill’ or ‘break’ (in the usual construction)
 - P is the patient argument of such a verb
 - S is the argument of a change-of-state verb like ‘fall’ or ‘die’
- (Haspelmath 2011a)

These comparative concepts can be applied to all languages **uniformly**, because all languages have verbs with these meanings and have nominal arguments with particular coding properties.

One might also use more **fine-grained semantic roles**, as in Hartmann et al. (2014):

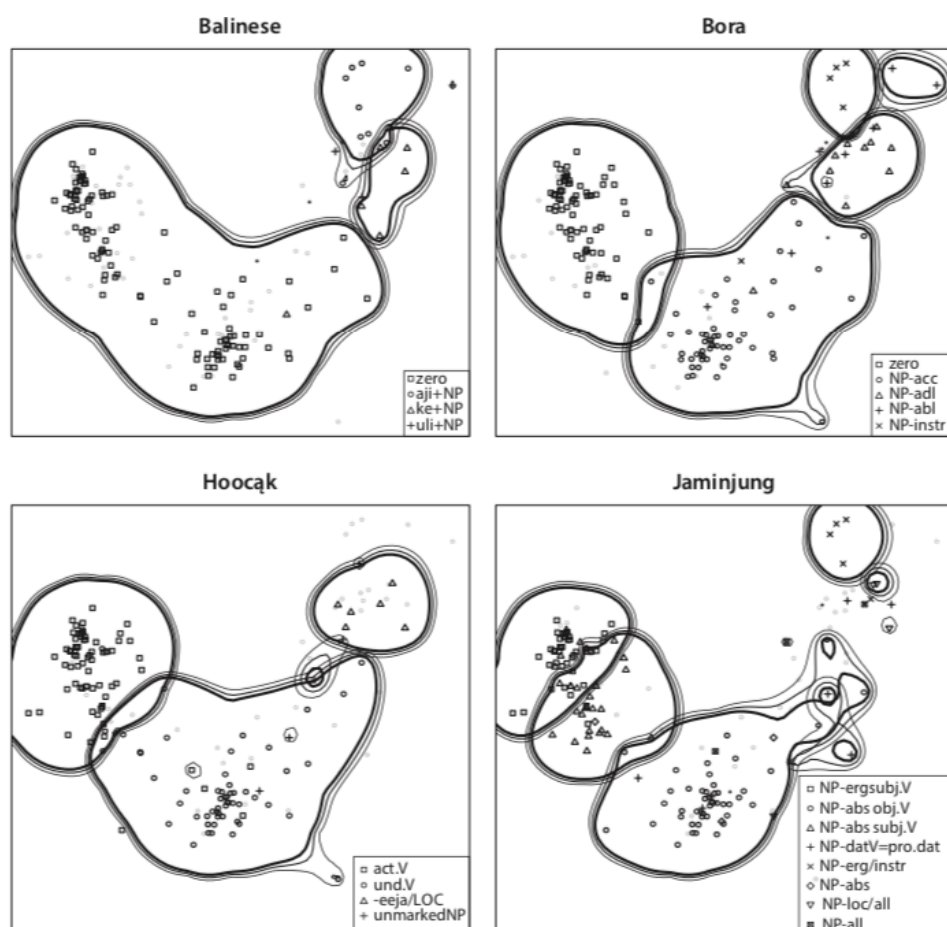


Figure 5. Four additional languages showing different distributional ranges of coding sets

Many authors define “subject” in such a way that **different criteria** are applied in **different languages**:

Mel’čuk (2013: x):

“...the SyntSubj [is] cross-linguistically universal. However, in a different sense, the SyntSubj is language-specific in so far as syntactic privileges are different in different languages: thus, in many Indo-European languages the main privilege of a clausal element is to impose agreement on the Main Verb, while in Malagasy it is to occupy the clause-final position.”

Riesberg et al. (2019: 524)

“We use the term ‘subject’ here as equivalent to what is termed ‘privileged syntactic argument’ (PSA) in Van Valin (2005) and elsewhere. A PSA is defined as the syntactic element that controls coding properties such as agreement and that is the pivotal element in complex constructions such as relativization, NP deletion, control, and so forth.”

This procedure is unlikely to pick out uniform phenomena across languages – if different subjects may be recognized by different criteria, how do we know that they are all “subjects”?

“I know it when I see it?”

– no, we need strict definitions of comparative concepts
that allow rigorous testing

4. Mark Baker’s theory of “dependent case”

Baker (2015) claims that grammatical case-marking patterns in the world’s languages are usually determined by dependent-case rules as in (4).

- (4) a. High case in the clause is ergative.
b. Low case in the clause is accusative.
c. High case in VP is dative.
e. High case in NP is genitive.
f. Unmarked case is nominative-absolutive.

Cls:	NP	→	ERG
VP:	NP	→	DAT
VP:	NP	→	unmarked = ABS

Ali-di za-z ktab wuga-na.
Ali-ERG I-DAT book give-AOR
‘Ali gave me a book.’

There is no clear way to determine what is “high” and “low”, but Baker is not worried by this:

“We should not expect to find many differences between c-command and a notion like thematic prominence, since the two notions are closely related.” (p. 81)

But elsewhere, he appeals to a movement operation in order to explain some case-marking effects. Thus, Sakha has differential object marking on definite nominals (p. 125-126):

- (22) a. Masha salamaat-y sie-te.
 Masha porridge-ACC eat-PAST.3sS
 'Masha ate the porridge.'

- b. Masha salamaat sie-te.
 Masha porridge eat-PAST.3sS
 'Masha ate porridge.'

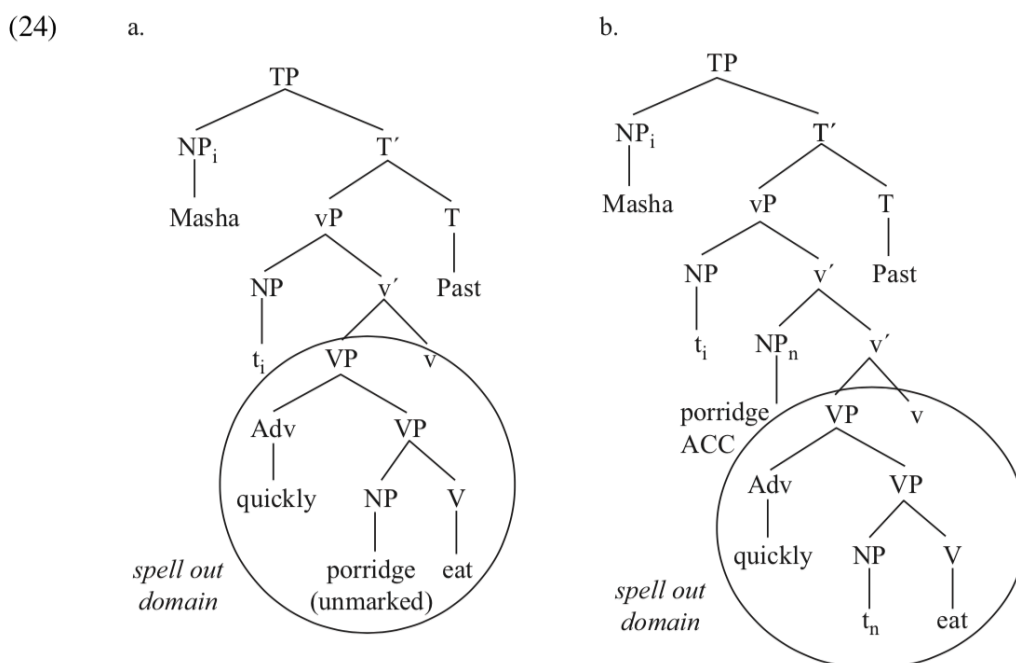
- (23) a. Masha salamaat-y türgennik sie-te.
 Masha porridge-ACC quickly eat-PAST.3sS
 'Masha ate the porridge quickly.'

(Object is * without ACC)

- b. Masha türgennik salamaat sie-te.
 Masha quickly porridge eat-PAST.3sS
 'Masha ate porridge quickly.'

(ACC would imply that the object is focused)

Baker's actual proposal for the movement is rather intricate and thus quite "deep" (p. 126):



But he misses the generalization that differential object marking is found on definite objects also when there is no word order alternation (as in Hebrew).

(How do we know that Hebrew indeed represents "the same phenomenon"? Because it obeys the same generalization, as can be determined by **objective hypothesis-testing**.)

If comparison is based on "deep analyses", with different considerations playing a role for different languages, then this means that we have to rely on the analysis

being correct. But if the analysis in turn relies on the general theory, this becomes a very intricate process...

Baker (2015: 222) on case of predicate nominals:

- NP
NP
- (5)
- | | | | | |
|-----------|-------------|----------------|--------------|------------|
| <i>Zi</i> | <i>buba</i> | <i>Joxanes</i> | <i>kešiš</i> | <i>ja.</i> |
| my | father(ABS) | Johannes(ABS) | priest(ABS) | COP |
- ‘My father Johannes is a minister.’ (from Haspelmath 1993: 311)

The subject nominal is “higher” than the predicate nominal – why doesn’t it bear Ergative case in Lezgian?

NP
EP
NP

This “E” (which projects an EP that prevents case assignment) is generally empty, but maybe attested in Tamil (affix *-aa* following the predicate nominal).

Baker’s methodological move:

Hypothesize that an underlying element exists, and look for phenomena that might match this phenomenon. If there is something to be found (whatever it may be), this is taken as conformation.

This was called “**diagnostic-fishing**” in an earlier paper (Haspelmath 2018: 102).

But **confirmation bias** is one of the best-known problems in all domains of knowledge acquisition. If there is no objective method of **measuring** and counting effects, one can easily be led astray.

Baker on PPs

Baker (2015: 185-186) observes that neither PPs nor oblique-marked nominals trigger ergative case on the subject:

- c. Ahwmad suona ulul laatt.
 Ahmed.ABS I.DAT next.to stand-PRES
 ‘Ahmed is standing next to me.’
 (Ingush, Nichols [2011: 401])
- a. Jose-(*kan) ochiti-ki raket-ai. (Shipibo)
 José-(*ERG) dog-DAT fear-IMPF
 ‘José fears the dog.’

But his theory predicts that *Jose* should be ergative because the object nominal is not a PP. So he hypothesizes that these are in fact PPs, with an empty adposition that assigns oblique case.

Baker posits not only empty adpositions (in PP) that assign case, but also surface adpositions that are really case forms (in NP), and surface case forms that are really adpositions (in PP) (p. 2, p. 9).

For such cases, he says that “the theory will have to decide (p. 13)”
(i.e. the theory is not motivated here).

But one can “hope that one can find some fine-grained syntactic properties which distinguish the two kinds [...] : a process of clefting, perhaps, or quantifier floating – the sorts of syntactic phenomena known to apply to NPs but not to PPs in some languages”

How far can the “depth of analysis” go?

Baker (2015: 25-26):

“We could contemplate taking a more radical view, which would claim that all languages are really tripartite languages in their syntactic case marking....

Some strict minimalist theorists might find this view attractive. However, I believe that this universalist view goes too far. Rather, I claim that languages are parameterized...”

But how can we tell whether Baker is right or wrong?

My proposal:

Let’s work with observations at the level of **community norms**, not with deeper analyses. Thus, any pedagogical grammar should be sufficient for comparative purposes.

5. Mel’čuk and Lezgian again

Mel’čuk provides the following definition of “subject” (1988: 163):

“A **grammatical subject** (GS) in **L** is either a basic GS or any other nominal that is most similar to the basic GS from the viewpoint of relevant syntactic properties {P_i}.”

And he observes:

“The syntactically privileged status of NPs is **language-specific**. Thus a syntactic property relevant in one language may turn out nonexistent or immaterial in another. For example, controlling the agreement of the main verb is a mark of privileged status in Russian or Italian, while it does not appear at all in Dyirbal, Lezgian or Japanese...”

He walks through a number of subject criteria and argues that the presumed Lezgian “ergative construction” lacks them (see Haspelmath 1991 for detailed discussion).

- (1) Алиди кицІ кьена
 Ali-di *kic’* *q’e-na*.
 Ali-ERG dog kill-AOR
 ‘Ali killed the dog.’

In a footnote (p. 244), he admits that A. E. Kibrik (1979-1980) has noted some phenomena that seem to contradict his analysis, because the ergative subject (and also the dative subject) seem to have the same privileges of occurrence in ‘want’ complements as the absolutive subject:

- (6) a. гададиз рушаз килигиз кІанзава
Gadadi-z [*riša-z* *kʰlig-iz* *kan-zava*].
 boy-DAT ABS girl-DAT look-INF want-IMPF
 ‘The boy wants to look at the girl.’
- b. гададиз руш катаз кІанзава
Gadadi-z [*riš* *kʰatʰa-z* *kan-zava*].
 boy-DAT ERG girl(ABS) beat-INF want-IMPF
 ‘The boy wants to beat the girl.’
- c. гададиз руш акваз кІанзава
Gadadi-z [*riš* *akʰwa-z* *kan-zava*].
 boy-DAT DAT girl(ABS) see-INF want-IMPF
 ‘The boy wants to see the girl.’
- d. гададиз руш вичиз килигна кІанзава
Gadadi-z [*riš* *üčʰü-z* *kʰlig-na* *kan-zava*].
 boy-DAT girl(ABS) self-DAT look-CVB want-IMPF
 ‘The boy wants the girl to look at him.’

Mel’čuk (1988: 247) says:

“I would still say that in Lezgian, the [absolutive nominal] of all verbs is the grammatical subject, that is, I would stand my ground anyway. The reason is that, to me, in Lezgian the property of PRO-control is motivated semantically... regardless of its actual surface-syntactic role... I consider the property “being a PRO-controller” as less weighty or less relevant than the previously considered subjecthood properties...”

But again, this represents arbitrary choice of criteria.

Unfortunately, this is not uncommon linguistics, though it is rarely admitted so clearly. Cf. also:

Börjars (1998: 44): “The behaviour of elements is often not totally consistent. This means that in order to arrive at the conclusion that an element is either a clitic or an affix, certain criteria must be assumed to be less crucial.” (cited by Haspelmath 2011b: 59)

6. What “depth of analysis” is good for

I conclude that the much-praised “detailed” and “deep” analyses of generative grammar seem to have little value for comparative purposes.

But of course, **“detailed” study is good** – we want to know everything about all languages. Not only all the words, but also all the constructions and all their interactions need to be studied.

Grammatical descriptions should be complete, and **consist of many volumes**.

But “deep” analysis is not necessarily good, because it is often unclear whether the “deep” generalizations are true.

In Haspelmath (1991; 1993), I ended up saying that there is some evidence for a Subject category in Lezgian, because this allows us to generalize over intransitive Absolutive, transitive Ergative, and experiencer Dative arguments, which behave alike in ‘want’ complements clauses (see 6a-d above).

But the evidence for this is hardly compelling.

7. Measurement uniformity and building-block uniformity

Instead of “depth of analysis”, comparative research is best based on objectively defined comparative concepts.

e.g. **transitive construction**:

= a minimal-clause construction with an A-argument and a P-argument (Haspelmath 2011a)

A-argument and P-argument:

= the ‘killer/breaker’ argument of the verbs ‘kill/break’, and arguments that have the same kind of argument coding

argument coding:

= flagging and indexing (cf. Haspelmath 2019)

ergative flag:

= a flag that can be used for the A but not for the P or the S

S-argument:

= the ‘faller/dieer’ argument of the verbs ‘fall/die’ and arguments that have the same kind of argument coding

On the basis of the notion of “ergative flag”, one can formulate and test universals about **differential ergative marking**, as seen in Godoberi (2a-c above). As we saw, these universals have a good functional explanation (in terms of efficient grammatical coding).

Baker (2015) operates in terms of invisible “case features”, which are often (but need not be) spelled out on the “surface”. Thus, he assumes zero spellout for the classical Australian split-ergative languages,

e.g. Diyari (Austin 1981)

	1st/2nd	other pronouns		full nouns	
A	Ø	-li		-li	(ERG)
P	- <u>na</u>	- <u>na</u>	(ACC)	-Ø	(ABS)
S	Ø	Ø		-Ø	(ABS)

Baker regards such languages as having “tripartite case”, with widespread zero-realization.

He thus misses the generalization that these languages conform to the high-level generalization about differential flagging:

(3) **The role-reference association universal** (Haspelmath 2020b)

Deviations from usual associations of role rank and referential prominence tend to be coded by longer grammatical forms if the coding is differential.

Arkadiev (2017): When a language has two different ergative markers, there is a (slight) preference for the longer ergative marker to be

Like all generative grammarians, Baker assumes **building-block uniformity**: All languages are basically made from the same innate building blocks.

These building blocks are both the stuff of language-particular analyses, **and** the basis of comparison.

But there are no arguments – building-block uniformity is simply assumed by generative grammarians, as if no alternative existed.

The better alternative is to have two different sets of concepts:

- descriptive categories for language-particular generalizations
 - comparative concepts for **measurement uniformity**
- (Haspelmath 2010; 2018)

Language-particular research creates **language-particular theories**.

Comparative research creates **comparative and general theories**.

The difference between generative (Baker-style) comparative grammar and Greenberg-style comparative grammar is that Baker-style comparison relies on innate building blocks, while Greenberg-style comparison does not.

It's a difference in **methodological choices (not ideologies)**.

How do we find out? We look which approach finds the better (more stable) theories which fit the world's linguistic diversity better – though we should probably invest less in building-block uniformity approach, because of its inherent unlikelihood.

8. Conclusion

The primary contrast is not between “typology” (Greenberg-style) and “parameters” (Baker-style), but between two methodological orientations:

- **measurement uniformity** (cf. Mel'čuk's emphasis on definitions)
- **building block uniformity** (cf. Cinque's cartographic ideas)

In chemistry, the building-block uniformity approach has worked well (cf. Mendeleyev's Periodic Table of Elements). It is conceivable that it will work in linguistics as well – I call it the **Mendeleyevian Vision** (cf. Baker 2001).

But measurement uniformity is the more tractable approach, which allows us to engage in systematic and quantitative language comparison.

In view of “Darwin’s Problem” (the low likelihood of a rich innate grammar evolving over a very short period), this approach is also more likely, and is the only one compatible with Berwick & Chomsky (2016).

References

- Arkadiev, Peter. 2017. Multiple ergatives. *Studies in Language* 41(3). 717–780. doi:[10.1075/sl.41.3.06ark](https://doi.org/10.1075/sl.41.3.06ark).
- Austin, Peter K. 1981. *A grammar of Diyari, South Australia*. Cambridge: Cambridge University Press.
- Baker, Mark C. 2001. *The atoms of language*. New York: Basic Books.
- Baker, Mark C. 2015. *Case: Its principles and parameters*. Cambridge: Cambridge University Press.
- Barrett, Lisa Feldman. 2006. Are emotions natural kinds? *Perspectives on Psychological Science* 1(1). 28–58. doi:[10.1111/j.1745-6916.2006.00003.x](https://doi.org/10.1111/j.1745-6916.2006.00003.x).
- Berwick, Robert C. & Noam Chomsky. 2016. *Why only us: Language and evolution*. Cambridge MA: MIT Press.
- Börjars, Kersti. 1998. *Feature distribution in Swedish noun phrases*. Oxford: Blackwell.
- Comrie, Bernard. 1978. Ergativity. In Winfred P. Lehmann (ed.), *Syntactic typology: Studies in the phenomenology of language*, 329–394. Austin: University of Texas Press.
- DeLancey, Scott. 1981. An interpretation of split ergativity and related patterns. *Language* 57. 626–657.
- Hartmann, Iren, Martin Haspelmath & Michael Cysouw. 2014. Identifying semantic role clusters and alignment types via microrole coexpression tendencies. *Studies in Language* 38(3). 463–484.
- Haspelmath, Martin. 1991. On the question of deep ergativity: The evidence from Lezgian. 44/45(1–2). 5–27. doi:[10.5281/zenodo.225289](https://doi.org/10.5281/zenodo.225289).
- Haspelmath, Martin. 1993. *A grammar of Lezgian* (Mouton Grammar Library 9). Berlin: Mouton de Gruyter.
- Haspelmath, Martin. 2007. Ditransitive alignment splits and inverse alignment. *Functions of Language* 14(1). 79–102.
- Haspelmath, Martin. 2010. Comparative concepts and descriptive categories in crosslinguistic studies. *Language* 86(3). 663–687. doi:[10.1353/lan.2010.0021](https://doi.org/10.1353/lan.2010.0021).
- Haspelmath, Martin. 2011a. On S, A, P, T, and R as comparative concepts for alignment typology. *Linguistic Typology* 15(3). 535–567.
- Haspelmath, Martin. 2011b. The indeterminacy of word segmentation and the nature of morphology and syntax. *Folia Linguistica* 45(1). 31–80.
- Haspelmath, Martin. 2018. How comparative concepts and descriptive linguistic categories are different. In Daniël Van Olmen, Tanja Mortelmans & Frank Brisard (eds.), *Aspects of linguistic variation: Studies in honor of Johan van der Auwera*, 83–113. Berlin: De Gruyter Mouton. <http://doi.org/10.5281/zenodo.570000>.
- Haspelmath, Martin. 2019. Indexing and flagging, and head and dependent marking. *Te Reo* 62(1). 93–115. doi:[10.17617/2.3168042](https://doi.org/10.17617/2.3168042).
- 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>.
- Kibrik, Alexandr (ed.). 1996. *Godoberi*. München: Lincom Europa.
- Lazard, Gilbert. 2002. Transitivity revisited as an example of a more strict approach in typological research. *Folia Linguistica* 36(3–4). 141–190.
- Mel'čuk, Igor A. 1981. Grammatical subject and the problem of the ergative construction in Lezgian. In Bernard Comrie (ed.), *Studies in the Languages of the USSR*, 229–276. Carbondale & Edmonton: Linguistic Research.
- Mel'čuk, Igor A. 1982. *Towards a language of linguistics: A system of formal notions for theoretical morphology*. Munich: Fink.
- Mel'čuk, Igor A. 1988. *Dependency syntax: Theory and practice*. Albany: State University Press of New York.

- Mel'čuk, Igor A. 2013. Syntactic subject, once again. In Valentina Apresjan, Boris Iomdin & Ekaterina Ageeva (eds.), *Proceedings of the 6th International Conference on Meaning-Text Theory, Prague, August 30-31, 2013*, iii–xxxiii. Prague.
- Nichols, Johanna. 2011. *Ingush grammar* (University of California Publications in Linguistics, 141). Berkeley: University of California Press.
- Polinsky, Maria. 2011. Linguistic typology and formal grammar. In Jae Jung Song (ed.), *The Oxford handbook of linguistic typology*. Oxford: Oxford University Press.
- Polinsky, Maria & Robert Kluender. 2007. Linguistic typology and theory construction: Common challenges ahead. *Linguistic Typology* 11(1). 273–283. doi:[10.1515/LINGTY.2007.022](https://doi.org/10.1515/LINGTY.2007.022).
- Riesberg, Sonja, Kurt Malcher & Nikolaus P. Himmelmann. 2019. How universal is agent-first? Evidence from symmetrical voice languages. *Language* 95(3). 523–561. doi:[10.1353/lan.2019.0055](https://doi.org/10.1353/lan.2019.0055).
- Roberts, Ian. 2019. *Parameter hierarchies and universal grammar*. Oxford: Oxford University Press.
- Schmidtke-Bode, Karsten & Natalia Levshina. 2018. Assessing scale effects on differential case marking: Methodological, conceptual and theoretical issues in the quest for a universal. In Ilja A. Seržant & Alena Witzlack-Makarevich (eds.), *Diachrony of differential argument marking* (Studies in Diversity Linguistics 19), 463–489. Berlin: Language Science Press.

Alignment and differential coding:

2B. Differential flagging

MARTIN HASPELMATH

Max Planck Institute for Evolutionary Anthropology

1. Argument coding splits

(Haspelmath 2021b)

- In many languages, the coding of core arguments depends on their REFERENTIAL PROMINENCE in one way or another

differential object marking is a well-known phenomenon, e.g.

- (1) Spanish (García García 2007) (animacy-based)

- a. *Conozco *(a) este actor.*
‘I know this actor.’
- b. *Conozco (*a) esta película.*
‘I know this film.’

- (2) Punjabi (Bhatia 1993: 172–174) (specificity-based)

- a. *Kataab vekho.*
book look.IMP.2PL
‘Look at a book.’
- b. *Kataab nūū vekho.*
book ACC look.IMP.2PL
‘Look at the book.’

(note: differential marking \simeq split marking)

- **prominence-conditioned splits** are cross-linguistically regular in a way that is surprising but apparently robust (Silverstein (1976), Comrie (1978), Moravcsik (1978a, 1978b), Dixon (1979), Bossong (1985; 1998) Lazard (2001))

- referential prominence has become widely known by names such as “**animacy hierarchy**” or “**empathy hierarchy**” (since the 1980s)

- the (a) example below shows **overt coding** of an argument, and the (b) example shows **zero coding** of the same argument

- (1) split subject (A) marking: Kham (Watters 2002)

- a. *no-ra-e zihm jə-ke-rə*
he-PL-ERG house.ABS make-PFV-3PL
‘They made a house.’

- b. *ŋa:-Ø zihm ŋa-jəi-ke*
I-NOM house.ABS 1SG-make-PFV
‘I made a house.’

(2) split object (P) marking: Sakha (Baker 2015: 4-5)

- a. *Masha salamaat-**y** türgennik sie-te.*
 Masha porridge-ACC quickly eat-PST.3SG.SBJ
 ‘Masha ate the porridge quickly.’
- b. *Masha türgennik salamaat-Ø sie-te.*
 Masha quickly porridge-Ø eat-PST.3SG.SBJ
 ‘Masha ate porridge quickly.’

In addition, argument coding splits may depend on the referential prominence properties of the SCENARIO:

- in monotransitive constructions, A and P
- in ditransitive constructions, R and T

(3) split marking of P only if A is third person: **Teop** (Mosel 2007: 10)

- a. (3 > 3)
*A beiko tenaa paa asun=u **ben-e** guu.*
 art child my TAM kill=imm OBJ-ART pig
 ‘My child has killed the pig.’
- b. (1 > 3)
Enaa paa dee ma=u e guu.
 1SG TAM carry DIR=IMM ART pig
 ‘I have brought a pig.’

(4) split marking of R obligatory if T is a personal pronoun, and R is a full nominal:
English

- a. (N > N)
*She gave Kim the money. (≈ She gave the money **to** Kim.)*
- b. (pers > pers)
*She gave him it. (≈ She gave it **to** him.)*
- c. (N > pers)
**She gave Kim it.*
- d. (N > pers)
*She gave it **to** Kim.*

All these are special cases of the high-level generalization in (5).

(5) **Universal 4: The role-reference association universal**

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

• proposed explanation of this universal:

a special case of the **form-frequency correspondence universal** of Haspelmath (2021a)

• The basic idea is that additional coding such as the accusative marker in (2a) or the dative marker in (4d) is required when it is least predictable and hence needed the most, i.e. that argument coding splits reflect a functional motivation.

Across a range of different situations,

– arguments with a **high-ranked role** (the transitive A-argument and the ditransitive R-argument) are **referentially prominent** in the usual case (or in other words, the most frequent case), and

– arguments with a **low-ranked role** (the transitive P-argument and the ditransitive T-argument) usually exhibit **lower referential prominence**.

Special coding by longer forms is used when a construction deviates from these usual associations.

2. Role rank and referential prominence

(6) usual role-reference association

Arguments with higher-ranked roles tend to be more referentially prominent, and vice versa.

This tendency is not meant as an abstract aspect of language structures (like the related notion of “harmonic alignment” of Aissen 2003), but as a concrete claim about discourse frequencies.

Role rank is defined very simply as in (7) (see Haspelmath 2011; 2015 for a precise definition of the comparative role-types A, P, T and R).

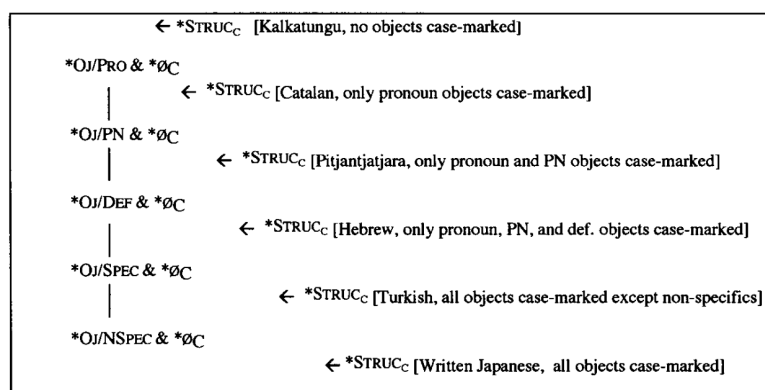
(7) role rank

In monotransitive constructions, the A is ranked higher than the P.

In ditransitive constructions, the R is ranked higher than the T.

It is of course an interesting question to ask why A should pattern with R, and why P should pattern with T, but for the claims I make here, it makes no difference, so the notion of role rank is not really important.

The idea of **usual association** of a role scale with referential scales is particularly clear in Aissen (2003), and I acknowledge the influence of her clear formulations (also in Aissen 1999). However, my proposed explanation of the generalizations is completely different and does not rely on a general notion of „role rank“.



Referential prominence is defined by the scales of INHERENT PROMINENCE and DISCOURSE PROMINENCE in (8).

(8) scales of referential prominence

a. **inherent prominence**

person scale:	1st/2nd > 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

- one could ask what these different kinds of semantic and discourse-pragmatic notions have in common that justifies subsuming them under referential prominence (cf. Shibatani (2006): “discourse relevance”)

- The use of the the term *prominence* in connection with split argument coding seems to have been introduced by Aissen (1999; 2003), inspired by phonological terminology, but it is by now well-established (e.g. Malchukov 2008; Bornkessel-Schlesewsky & Schlewsky 2009; Lockwood & Macaulay 2012).

- It is also important to be aware that the notions in (7) and (8a-b) are intended as **comparative concepts** for cross-linguistic comparison, not as necessarily corresponding closely to **language-particular categories**. Categories differ across languages, but the concepts in (7) and (8) can be applied to all languages.

3. Usual associations and universals of coding splits: Two types

On the one hand, we can look at individual arguments:

(9) single-argument association tendencies

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

For example, clauses such as ‘The dog found a bone’ and ‘She gave the boy a key’, with definite A and R, and indefinite P and T, are more usual than ‘A rock hit the hiker’ or ‘She gave a boy the key’.

On the other hand, we can look at **both arguments** within a two-argument scenario simultaneously and compare their referential prominence values:

(10) scenario association tendencies

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

(13) **Universal 5: 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 and for non-prominent A/R arguments.

- an asymmetric split is a split in which one of the coding types involves **longer** coding. In the majority of cases, this means overt coding contrasting with zero coding, but we sometimes also see examples of shorter and longer overt coding.
- simple examples illustrating the universal are **overt accusatives** in more prominent P-arguments (e.g. animates vs. inanimates, cf. English *he/hi-m* vs. *it/it*), or **overt datives** for less prominent R-arguments (e.g. *give the money to a beggar* vs. *give the beggar money*).

The universal in (13) is a more general formulation that subsumes the famous split P flagging universal in (14). An illustration was given in (2).

(14) **Universal 6: differential P flagging** (\approx DOM)

If a language has an asymmetric P flagging split, then the flagging is longer for prominent P-arguments.

The more general formulation also includes split A flagging (“differential subject marking”), as seen in (1), as well as split R and T flagging (illustrated below...).

4. Single-argument splits in monotransitive constructions

4.1. Split P-flagging (= differential object marking)

As mentioned earlier, differential P-flagging, traditionally known as “differential object marking”, has been widely discussed in the earlier literature.

4.1.1. Animacy-conditioned split P-flagging

(17) Spanish (García García 2007)

- Conozco *(a)este actor.*
- Conozco (*a) esta película.*

4.1.2. Nominality-conditioned split P-flagging. Some languages have P-flagging only on personal pronouns, but not on full nominals. A well-known example of such a language is English (*he* vs. *him*, *she* vs. *her*, etc.).

4.1.3. Givenness-conditioned split P-flagging. It has been known since Thomson (1912) that split P-flagging is sometimes conditioned by givenness (or “topicality”) (see also Dalrymple & Nikolaeva (2011)).

(19) Persian (Dalrymple & Nikolaeva 2011: 108-112)

- man ketâb-râ xarid-am.*
I book-ACC buy.PST-1SG
‘I bought the book.’
- man sib-i(?*-râ) xord-am.*
I apple-INDEF(-ACC) eat.PST-1SG
‘I saw an apple.’

- c. *ki mašin-i-?*(râ) did?*
 who car-INDEF(ACC) see.PST[3SG]
 ‘Who saw a car?’

4.1.4. Person-conditioned split P-flagging. In Abruzzese (an Italo-Romance variety), special P-flagging by the preposition *a* occurs only with locuphoric (1st and 2nd) personal pronouns (D’Alessandro 2017: 8).

(20) Abruzzese (dialect of Arielli)

- a. *So vistə a mme/ a tte.*
 be.1SG seen to me/ to you
 ‘I have seen me/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 Mary/ to them/ to them

4.2. Split A-flagging

• split A-flagging (“differential subject marking”) is not often conditioned by animacy/specificity – but basically the same generalizations apply to ergative systems.

4.2.1. Person-conditioned split A-flagging. Systems on which ergative marking is restricted in that it does not occur on locuphoric person forms are found in Australia, South Asia, and in two families of the Caucasus (Kartvelian, Nakh-Daghestanian).

(21) Warrgamay

- a. *ngana-Ø gaga-ma*
 we-NOM go-FUT
 ‘We will go.’
- b. *ngana-Ø ngulmburu-Ø ngunda-lma*
 we-ERG woman-ACC see-FUT
 ‘We will see the woman.’
- c. *maal-du ngulmburu-Ø ngunda-lma*
 man-ERG woman-ACC see-FUT
 ‘The man will see the woman.’

4.2.2. Focus-conditioned split A-flagging: when the ergative marker occurs only when the A-argument is focused (cf. McGregor 2010)

(22) Central Tibetan (Tournadre 1995: 264)

- a. *khōng khāla’ so-kiyo:re’*
 he food make-IPFV.GNOM
 ‘He prepares the meals.’
- b. *khōng-ki’ khāla’ so-kiyo:re’*
 he-ERG food make-IPFV.GNOM
 ‘HE prepares the meals.’

5. Single-argument splits in ditransitive constructions (cf. Haspelmath 2007)

5.1. Split T coding

5.1.1. Specificity-conditioned split T-flagging

In Akan, another Kwa language from West Africa, the T argument must be indefinite in a simple double-object construction, and a construction with a special T-marking serial verb must be used instead (*dè* lit. 'take')

(25) Akan (Osam 1996: 63-64)

- a. *Kofi ma-a abofra no akokɔ*
Kofi give-COMPL child DEF chicken
'Kofi gave the child a chicken.'
- b. **Kofi ma-a abofra no akokɔ no*
Kofi give-COMPL child DEF chicken DEF
(‘Kofi gave the child the chicken.’)
- c. *Kofi de akokɔ no ma-a abofra no*
Kofi **take** chicken DEF give-COMPL child DEF
'Kofi gave the chicken to the child.'

5.2. Split R-flagging

5.2.1. Person-conditioned split R-marking

(26) French person clitics

	T (ACC)	R (DAT)		T (ACC)	R (DAT)
1SG	<i>me</i>	<i>me</i>	1PL	<i>nous</i>	<i>nous</i>
2SG	<i>te</i>	<i>te</i>	2PL	<i>vous</i>	<i>vous</i>
3SG	<i>le, la</i>	<i>lui</i>	3PL	<i>les</i>	<i>leur</i>

(27) Warlpiri (Hale 1972)

- a. *ngajulu-rlu ka-rna-ngku karli-Ø yi-nyi nyuntu-ku*
I-ERG PRS-1SG.NOM-2SG.ACC boomerang-ABS give-NPST you-DAT
'I am giving you a boomerang.'
- b. *ngajulu-rlu ka-rna-rla karli-Ø yi-nyi kurdu-ku*
I-ERG PRS-1SG.NOM-3.DAT boomerang-ABS give-NPST child-DAT
'I am giving a boomerang to the child.'

Table 2: Object person forms (1st person singular and 3rd person masculine singular)

language	1st person R/T '(to) me'	aliophoric R 'to him'	aliophoric T 'him'
French	<i>me</i>	<i>lui</i>	<i>le</i>
Tangale ^a	<i>-no/-no</i>	<i>-ni/-ni</i>	<i>mbéëndâm</i>
Yimas ^b	<i>ŋa-</i>	<i>-(n)akn</i>	<i>na-</i>
Krongo ^c	<i>àʔàŋ</i>	<i>àníŋ</i>	<i>ìʔìŋ</i>

^a Jungraithmayr (1991: 36) ^b Foley (1991) ^c Reh (1985)

5.2.2. Nominality-conditioned split R-flagging

(28) Northeastern Neo-Aramaic of Telkepe (Coghill 2010)

(full nominal R)

- a. *wəl-lə pāṛə 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 hadiynn*
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)

5.2.3. Animacy-conditioned split R-flagging. In Yakkha, a Kiranti language of Nepal, R-arguments are in the Locative case when inanimate, but otherwise in the (zero-coded) Absolutive case.

(29) Yakkha (Schackow 2012: 161-162)

- a. *ka nniŋda photo-ci ham-biʔ-meʔ-nenin=ha*
1SG[NOM] 2PL[NOM] photo-PL[NOM] distribute-BEN-NPST-1>2PL-PL
'I distribute the photos to you all.'
- b. *sarkar=ŋa yaŋ tenten=be ŋ-hapsu-bi-ci=ha*
government=ERG money[NOM] villages=LOC 3PL.A-distribute-BEN-3PL.P=PL
'The government distributed the money to the villages.'

5.2.4. Specificity-conditioned split R-flagging

In Wolof, an Atlantic language of Senegal, a dative flag is required on R when it is indefinite.

(30) Wolof (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.'

6. Some scenario splits

6.1. Monotransitive, person-conditioned. Yukaghir: accusative flag is required on P when the A is allophoric (like Teop, see above).

(31) Kolyma Yukaghir (Maslova 2003)

a. *met es'ie tet pulut-kele kudede-m*
 my father.NOM your husband-ACC kill-TR.3SG
 'My father has killed your husband.'

b. *met tolow kudede*
 I.NOM deer.NOM kill.TR.1SG
 'I killed a deer.'

6.2. Monotransitive, definiteness-conditioned. Eastern Khanty has Ergative case on the A-argument when the P-argument is specific (cf. Baker 2015: 128)

(33) Eastern Khanty

a. *Mä t'äkäjälämnä ula manyäləm.*
 we.DU.NOM younger.sister.COM berry pick.PST.1PL.SBJ
 'I went to pick berries with my younger sister.'

b. *Mə-ŋən ləyə əllə juy kanŋa aməyaloy.*
 we-ERG them large tree beside put.PST.3PL.OBJ/1PL.SBJ
 'We put them (pots of berries) beside a big tree.'

6.3. Ditransitive, conditioned by person of T

This is what is known in the literature as “person-case constraint” (PCC) (Haspelmath 2004).

Bulgarian: dative preposition is required on R when T is 1st/2nd person

(34) Bulgarian (Hauge 1976 [1999])

a. (3>3) *Az im ja preporâčvam.*
 I 3PL.DAT 3SG.F.ACC recommend.PRES.1SG
 'I recommend her to them.'

b. (3>2) **Az im te preporâčvam.*
 I 3PL.DAT 2SG.ACC recommend.PRES.1SG
 'I recommend you to them.'

c. *Az te preporâčvam na tjah.*
 I 2SG.ACC recommend.PRES.1SG to them
 'I recommend you to them.'

6.4. Ditransitive, conditioned by by nominality of T

In many varieties of English (especially American, it seems), the R cannot be coded in the simplest way when the T is a person form rather than a full nominal. In these varieties, (37a) is unacceptable (**Pat showed him it*).

- (36) a. (pers > nom, downstream)
Kim showed me his house.
- b. (nom > nom, balanced)
Lee showed her brother her new house.
- (37) a. (pers > pers, balanced)
**Pat showed him it.*
- b. OK: *Pat showed it to him.*
- c. (nom > pers, upstream)
**Pat showed his wife it.*
- d. OK: *Pat showed it to his wife.*

7. Explanation: Frequency-based coding efficiency

(57) Universal 4: The role-reference association universal

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

a functional-adaptive constraint:

- Languages tend to use **efficient coding**, i.e. zero or short coding for frequently occurring meanings and functions, and overt and long coding for rarely occurring functions.
- Through adaptability in language use, languages come to have or restore efficient patterns.

For role-reference associations, this kind of explanation has long been advocated, for example by Caldwell (1857: 276), who deserves to be quoted again here (cf. Filimonova 2005: 78):

“[. . .] the principle that it is more **natural** for rational beings to act than to be acted upon; and hence when they do **happen to be** acted upon – when the nouns by which they are denoted are to be taken objectively – it becomes necessary, in order to **avoid misapprehension**, to suffix to them the objective case-sign”

References

- Aissen, Judith. 1999. Markedness and subject choice in Optimality Theory. *Natural Language & Linguistic Theory* 17. 673–711.
- Aissen, Judith. 2003. Differential object marking: Iconicity vs. economy. *Natural Language & Linguistic Theory* 21(3). 435–483.
- Baker, Mark C. 2015. *Case*. Cambridge: Cambridge University Press.
- Becher, Jutta. 2005. Ditransitive Verben und ihre Objekte im Wolof: Positionsregeln und Kombinierbarkeit. *Hamburger afrikanistische Arbeitspapiere (HAAP)* 3. 13–27.
- Bickel, Balthasar. 1995. In the vestibule of meaning: Transitivity inversion as a morphological phenomenon. *Studies in Language* 19(1). 73–127. doi:10.1075/sl.19.1.04bic.
- Bossong, Georg. 1985. *Differenzielle Objektmarkierung in den neuiranischen Sprachen*. Tübingen: Narr.

- Bossong, Georg. 1998. Le marquage différentiel de l'objet dans les langues d'Europe. In Jack Feuillet (ed.), *Actance et valence dans les langues de l'Europe*, 193–258. Berlin: Mouton de Gruyter.
- Essegbey, James. 2010. Inherent complement verbs and the basic double object construction in Gbe. In Enoch Oladé Aboh & James Essegbey (eds.), *Topics in Kwa syntax*, 177–193. Dordrecht: Springer.
- Filimonova, Elena. 2005. The noun phrase hierarchy and relational marking: Problems and counterevidence. *Linguistic Typology* 9(1). 77–113.
- Georgi, Doreen. 2012. A local derivation of global case splits. In Artemis Alexiadou, Tibor Kiss & Gereon Müller (eds.), *Local modelling of non-local dependencies in syntax*, 305–336. (Linguistische Arbeiten). Tübingen: Niemeyer.
- Haspelmath, Martin. 2010. Framework-free grammatical theory. In Heine, Bernd & Narrog, Heiko (eds.), *The Oxford Handbook of Linguistic Analysis*, 341–365. Oxford: Oxford University Press.
- Haspelmath, Martin. 2011. On S, A, P, T, and R as comparative concepts for alignment typology. *Linguistic Typology* 15(3). 535–567.
- Haspelmath, Martin. 2015. Ditransitive constructions. *Annual Review of Linguistics* 1. 19–41.
- Haspelmath, Martin. 2019. Can cross-linguistic regularities be explained by constraints on change? In Schmidtke-Bode, Karsten & Levshina, Natalia & Michaelis, Susanne Maria & Seržant, Ilja A. (eds.), *Competing explanations in linguistic typology*, 1–23. Berlin: Language Science Press. (<http://langsci-press.org/catalog/book/220>)
- Haspelmath, Martin. 2020. The structural uniqueness of languages and the value of comparison for description. *Asian Languages and Linguistics* 1(2). 346–366. (doi:10.1075/alal.20032.has)
- Haspelmath, Martin. 2021a. Explaining grammatical coding asymmetries: Form-frequency correspondences and predictability. *Journal of Linguistics* 57(3). 605–633. (doi:10.1017/S0022226720000535)
- Haspelmath, Martin. 2021b. Role-reference associations and the explanation of argument coding splits. *Linguistics* 59(1). 123–174. (doi:10.1515/ling-2020-0252)
- Haspelmath, Martin. 2021c. General linguistics must be based on universals (or nonconventional aspects of language). *Theoretical Linguistics* 47(1–2). 1–31. (doi:10.1515/tl-2021-2002)
- Joseph, John E. 2022. Saussure's universal grammar, Chomsky's structuralism. In Cosenza, Giuseppe & Forel, Claire A. & Puskas, Genoveva (eds.), *Saussure and Chomsky: Converging and diverging*, 19–36. Berne: Peter Lang.
- Keine, Stefan. 2010. *Case and agreement from fringe to core: A minimalist approach*. Berlin: De Gruyter.
- Lockwood, Hunter T. & Monica Macaulay. 2012. Prominence hierarchies. *Language and Linguistics Compass* 6(7). 431–446. doi:10.1002/lnc3.345.
- Malchukov, Andrej. 2008. Animacy and asymmetries in differential case marking. *Lingua* 118. 203–221.
- Maslova, Elena. 2003. *A grammar of Kolyma Yukaghir*. Vol. 27. (Mouton Grammar Library). Berlin / New York: Mouton de Gruyter.
- Moravcsik, Edith A. 1978a. On the case marking of objects. In Joseph H Greenberg (ed.), *Universals of human language*, vol. 4: Syntax, 249–289. Stanford: Stanford University Press.
- Moravcsik, Edith A. 1978b. On the distribution of ergative and accusative patterns. *Lingua* 45. 233–279.
- Reinöhl, Uta. 2016. A single origin of Indo-European primary adpositions?: Unveiling the Indo-Aryan branch-off. *Diachronica* 33(1). 95–130. (doi:10.1075/dia.33.1.04rei)
- Rude, Noel. 2009. Transitivity in Sahaptin. *Northwest Journal of Linguistics* 3(3). 1–37.
- Silverstein, Michael. 1976. Hierarchy of features and ergativity. In Robert M. W. Dixon (ed.), *Grammatical categories in Australian languages*, 112–171. Canberra: Australian National University.
- Smirnova, Anastasia & Jackendoff, Ray. 2017. Case assignment and argument realization in nominals. *Language* 93(4). 877–911. (doi:10.1353/lan.0.0225)
- Stroński, Krzysztof. 2009. Variation of ergativity patterns in Indo-Aryan. *Poznań Studies in Contemporary Linguistics* 46(2). 237–253. doi:10.2478/v10010-010-0012-8.
- Watters, David E. 2002. *A grammar of Kham*. Cambridge: Cambridge University Press.
- Whitney, William D. 1875. *The life and growth of language*. London: Henry S. King & Co.
- Zúñiga, Fernando. 2006. *Deixis and alignment: Inverse systems in indigenous languages of the Americas*. (Typological Studies in Language). Amsterdam: Benjamins.

Curso extracurricular: Comparing grammars worldwide (UNAM, 20-24 October 2025)

Alignment and differential coding: 2C. Argument indexing vs. argument flagging

MARTIN HASPELMATH

Max Planck Institute for Evolutionary Anthropology

1. Differential object flagging and differential object indexing

let us compare:

- differential object flagging (DOF)
- differential object indexing (DOI)
- the two corresponding types of differential subject marking (DSF and DSI)

Differential object marking (DOM) has long been prominent in Romance linguistics and beyond (e.g. Bossong 1991; Aissen 2003; Kabatek et al. 2021):

(1) Spanish (Fábregas 2013: 1)

a. *Encontré un problema.*

I.found a problem

‘I found a problem.’

b. *Encontré **a** un superviviente.*

I.found **ACC** a survivor

‘I found a survivor.’

DOM describes a situation in which two classes of patient objects are marked in **different ways** depending on their intrinsic properties (such as animacy) or on their discourse properties (such as definiteness).

Moravcsik (1978) and Bossong (1985): called the phenomenon
differential object marking

a more precise term for (1a-b): ***differential object flagging (DOF)***

(*flag* is a cover term for case and adpositional marking; Haspelmath 2019a)

Crucially: DOM comes in two subtypes:

DOF (differential object flagging)

DOI (***differential object indexing***)

the variable use of person indexing (or “agreement”, or “clitic doubling”),
as in (2)-(3)

(2) Romanian (Irimia 2020: 429)

a. *Maria a văzut o stea.*
 Maria has seen a star
 ‘Maria saw a star.’

b. *Maria l-a văzut pe Ion.*
 Maria **3SG.OBJ**-has seen **ACC** Ion
 ‘Maria saw Ion.’ (lit. ‘Maria saw him, John’)

(3) Mauwake (Trans-New Guinea; Berghäll 2015: 97; 216)

a. *Bom=iya kateres=iya fuurk-a-mik.*
 bomb=COM cartridge=COM drop-PST-3PL
 ‘They dropped bombs and cartridges.’

b. *Emeria wia=amukar-e-k.*
 woman **3PL.ACC**=scold-PST-3SG
 ‘She scolded the women.’ (lit. ‘She scolded them, the women.’)

Romanian: both differential object flagging (with the preposition *pe*) and differential object indexing (prefix *l-*) in the same sentence

Are the two phenomena (DOF and DOI) “the same” at some level? This has often been thought:

(4) **Presumption of DOM unity**

Differential object flagging and differential object indexing are manifestations of the same underlying or motivating factor.

e.g. Dalrymple & Nikolaeva (2011: 1-2):

“we understand DOM as covering both agreement and case-marking (case or adpositional marking on the object)”

e.g. Kalin (2018: 113):

“On an inclusive conception of DOM, which I adopt here, marking may take the form of case (e.g., Persian, Turkish), an adposition (e.g., Spanish, Hindi), agreement (e.g., Swahili, Ostyak), or clitic doubling (e.g., Macedonian, Catalan).”

However:

- there are good reasons to reject the presumption of unity of DOM
- its two subtypes DOF and DOI should probably be explained in different ways (see also Just 2024)

The two central observations:

DOI and DSI is **uniformly** driven by **referential prominence**,
 but DOF and DSF show mirror-image patterns:

DOF is found when the object is referentially prominent
 DSF is found when the subject is referentially **unprominent**

The latter observation can be explained in terms of the **efficiency theory of asymmetric coding** (Haspelmath 2021a; 2021b).

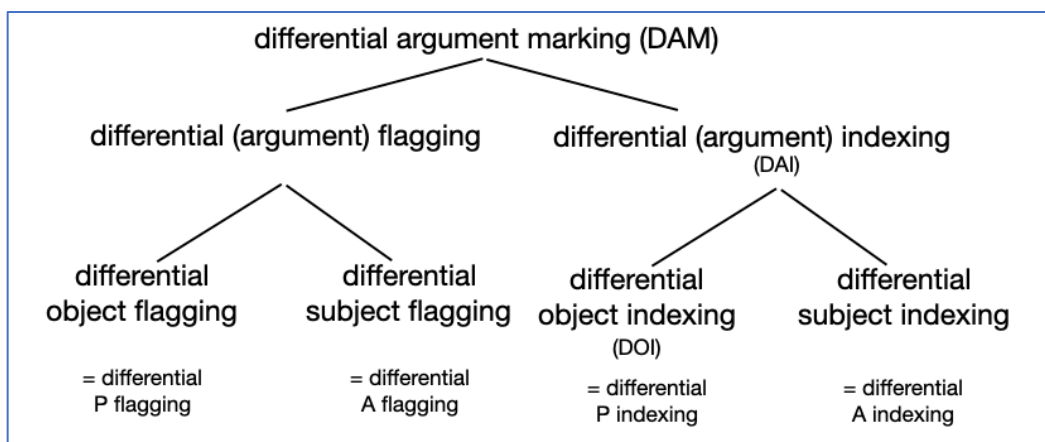


Figure 1. The terminology of differential argument marking

2. Differential object indexing (DOI) and the prominent indexing universal

2.1. Referential prominence

Differential argument indexing is typically conditioned by *referential prominence*, e.g.

Romanian: indexing of definite objects but not of indefinite inanimate objects
 Mauwake: indexing of animate objects but not inanimate objects.

This is a very general pattern

(cf. Morimoto 2022; Iemmolo 2011; Haig 2018; Just 2022; 2023; 2024):

(5) The prominent indexing universal (first version)

When argument indexing is differential and is conditioned by referential prominence, all languages prefer indexing of referentially prominent arguments.

Note: *argument indexing* = person-number indexing (Haspelmath 2013),
 i.e. argument marking by bound person-number forms

no distinction between “agreement” and “pronouns”,
 or between marking by “affixes” or “clitics” !!

object indexing is typically known as “clitic doubling” in Romance languages
 (e.g. Anagnostopoulou 2006; Poletto & Pinzin 2024).

six scales of referential prominence (or topicworthiness, or salience) (e.g. Aissen 1999; Siewierska 2004: 149; Lockwood & Macaulay 2012; Haspelmath 2021b: 127-128):

(6) scales of inherent prominence

person scale:	1st/2nd > 3rd person
full nominality scale:	person form (independent or index) > full nominal
animacy scale:	human (> animal) > inanimate

(7) scales of discourse prominence

specificity scale:	definite (> specific indefinite) > indefinite nonspecific
givenness scale:	discourse-given > discourse-new
focus scale:	background > focus

[Incidentally: these are comparative scales, not descriptive scales,
cf. Haspelmath 2015]

The prominent indexing universal means that if an argument is higher on one of these scales, it is more likely to be indexed than if it lower.

2.2. Pronominality-conditioned DOI

In some languages such as Spanish, object indexing occurs with personal pronoun objects but not with full nominals.

(8) Spanish

- a. ***Lo** veo a él.*
3SG.OBJ I.see ACC him
 'I see him.'

- b. *Veo a Pedro.*
 I.see ACC Pedro
 'I see Pedro.'

(9) French (Kayne 2000: 165)

- a. *Jean **la** connaît elle.*
 Jean **her** knows her
 'Jean knows HER.'

- b. **Jean la connaît Marie.*
 Jean her knows Marie
 'Jean knows Marie.'

Instead of "languages preferring" A over B, we can state the universal (more precisely) in implicational terms:

(10) **The prominent indexing universal** (second version)

If a referentially less prominent argument is indexed in a language,
the argument types higher on the scale are also indexed.

2.3. Animacy-conditioned DOI

In some languages, object indexing occurs with animate but not with inanimate objects.

(11) Oneida (Iroquoian; Koenig & Michelson 2015: 6)

a. *khe-nú-weh-se?*
1SG>3-like-HAB
 ‘I like her (or them).’

b. *k-nú-weh-se?*
 1SG.A-like-HAB
 ‘I like it.’

(12) Bunaq (Timor-Alor-Pantar; Schapper 2020: 125)

a. *Markus zo poi.*
 Markus mango choose
 ‘Markus chose a mango.’

b. *Markus zap go-poi.*
 Markus dog **3.OBJ**-choose
 ‘Markus chose a dog.’

2.4. Definiteness-conditioned DOI

In some languages, object indexing occurs with definite objects but not with indefinite objects.

(14) Bulgarian (cf. Compensis 2022: 30)

a. *Kuče-to goni kotka.*
 dog-DEF.N chases cat-DEF.F
 ‘The dog chases a cat.’

b. *Kuče-to ja goni kotka-ta.*
 dog-DEF.N **3SG.F.OBJ** chases cat-DEF.F
 ‘The dog chases the cat.’

2.5. Topicality-conditioned DOI

In some languages, object indexing occurs with topical objects but not with nontopical objects.

(15) Sanvalentinense (Abruzzo) (D’Alessandro 2017: 7)

a. *Ajə cciosə li pelliṣṣə.*
 I.have killed.M.SG the chickens.M.PL
 ‘I have been killing chickens.’

b. *Ajə ccisə li pelliṣṣə.*
 I.have **killed.M.PL** the chickens.M.PL
 ‘I have killed the chickens.’

(16) Tundra Nenets (Dalrymple & Nikolaeva 2011: 128)

- a. ('Whom did Vanya hit?')
Wanya Pyetya-m ladə°. (**ladə°da*)
 Vanya Petya-ACC hit.3SG.SBJ
 'Vanya hit PETYA.'
- b. ('What did a/the man do to a/the reindeer?')
Xasawa ti-m xada°-da. (**xada°*)
 man reindeer-ACC killed.3SG.SBJ-**3SG.OBJ**
 'A/the man KILLED a/the reindeer.'

3. Differential subject indexing (DSI)

Languages may also exhibit differential subject marking (DSM)

(see de Hoop & de Swart 2009)

two subtypes: differential subject flagging (DSF)
 differential subject indexing (DSI)

DSI is typically conditioned by **referential prominence**, like DOI, e.g.

(17) Awing (Bantu; Fominyam & Georgi 2021: 84)

- a. *Alombah (a) nə n-náŋnə məʒíə əzoonə.*
 Alombah **3SG** PST N-cook food yesterday
 'Alombah cooked the food yesterday.' (topical subject)
- b. *Wə (*a) nə n-náŋnə məʒíə əzoonə.*
 who (3SG) PST N-cook food yesterday
 'Who cooked the food yesterday?'

3.1. DSI conditioned by inherent prominence

In Nez Perce, only **animate** nominals trigger number agreement (Deal 2016: 318). Again, this applies both to subject indexing and object indexing.

- (18) a. *Lep-ít pícpic hi-w-s-íix ʔiníit-pe.*
 two-SUF cat 3.SUBJ-be-PRES-**S.PL** house-LOC
 'Two cats are in the house.'
- b. *Lep-ít cepéepy'uxtin' hii-we-s ʔiníit-pe.*
 two-SUF pie 3.SUBJ-be-PRES house-LOC
 'Two pies are in the house.'

In Welsh, only **personal pronouns** (but not full nominals) trigger subject person marking:

(19) Welsh (Borsley et al. 2007: 199)

- a. *Gwelo-**n** nhw ddraig.*
saw-**3PL** they dragon.
'They saw a dragon.'
- b. *Gwel-odd y bechgyn ddraig.*
saw-3SG the boys dragon
'The boys saw a dragon.'

3.2. DSI conditioned by discourse prominence

in English and French, postverbal (i.e. nontopical) plural subjects do not (need to) occur with plural verb forms:

(20) English

- a. *The three women **are** IN THE ROOM.*
- b. *There's three WOMEN in the room.*

(21) French

- a. *Les trois femmes **sont** (*est) venues.*
the three women **are** (is) come:PP:FEM:PL.
'The three women CAME.'
- b. *Il **est** (*sont) venu trois femmes.*
it **is** (*are) come:PP:MASC:SG three women
'There came three WOMEN.'

NOTE:

While DOI is often seen as “**additional indexing**” (or “doubling”), differential subject indexing (DSI) tends to be taken as “**absent indexing**” (the presence of indexing is seen as the default).

(e.g. Lambrecht's (2000: §3.2): “suspended subject-verb agreement”)

3.4. Summary on differential indexing

- both differential object indexing (DOI) and differential subject indexing (DSI) obey the prominent indexing universal:

(10) **The prominent indexing universal** (second version)

If in a language, a referentially less prominent argument is indexed,
the argument types higher on the scale are also indexed.

(the evidence for this universal is only impressionistic, but it has been known for some time, and no substantial counterevidence seems to have come to light)

4. Differential object flagging (DOF) and differential subject flagging (DSF)

- another example of DOF:

(24) Zhoutun Mandarin (Zhou 2022: 117, 24)

- a. 你苹果吃了么?
Ni phĩkur tʃhi=lɔ mɿ?
 you apple eat=PFV Q
 ‘Did you eat apples?’
- b. 苹果哈我一个吃了。
Phĩkur=xa ŋɿ i=kɿ tʃhi=lɔ.
 apple=ACC I one=CLF eat=PFV
 ‘An apple I ate.’

- another example of differential subject flagging (DSF)

in Punjabi, **only full-nominal A-arguments** are ergative-flagged:

(27) Punjabi (Chandra & Kaur 2017: 26)

- a. *Munde-ne rottii khaaḍḍ-ii.*
 boy-ERG bread(F) eat-F.SG
 ‘The boy ate bread.’
- b. *Maĩ(*-ne) rottii khaaḍḍ-ii.*
 I(-ERG) bread(F) eat-F.SG
 ‘I ate bread.’

key observation:

The overall generalization for differential flagging of core arguments is **not uniform**:

While both DOF and DSF are typically conditioned by referential factors, the **directionality of the scalar generalizations is the opposite!**

(29) The core argument flagging universal

- (a) If in a language, a referentially **less prominent** P-argument is accusative-flagged, the argument types higher on the scale are also flagged.
- (b) If a referentially **more prominent** A-argument is ergative-flagged, the argument types lower on the scale are also flagged.

The pattern for P-arguments in (a) is thus the **mirror image** of the pattern A-arguments in (b) (as noted by Dixon 1994).

Moreover, as noted in Haspelmath (2021b), the generalization in (33) can be extended to the T- argument and the R-argument of ditransitive clauses:

The T-argument behaves like the P-argument in its flagging preferences, while the R-argument behaves like the A-argument.

e.g. English

- (30) a. *He gave his sister the money.*
 b. *??He gave a beggar the money.*
 c. *He gave the money **to** a beggar.* (differential recipient flagging)

Given the role ranking **A/R > P/T**, we can thus formulate the still more general flagging universal in (31).

(31) **The role-reference association universal** (Haspelmath 2021b)

Deviations from usual associations of role rank and referential prominence tend to be coded by longer grammatical forms if the coding is asymmetric.

5. Explaining differential flagging

Expectedness leads to short coding:

(34) **explanation in terms of expectation sensitivity**

Speakers add extra coding material when a meaning is unexpected in its context, and this may be grammaticalized.

An explanation along these lines has long been advocated. Some relevant quotations from the 20th century and from the less well-known literature are given below.

“[...], the most **natural kind of transitive construction** is one where the A is high in animacy and definiteness, and the P is lower in animacy and definiteness; and any **deviation** from this pattern leads to a more **marked** construction.” (Comrie 1989: 128)

(See Haspelmath 2021b)

6. Differential indexing is uniform across roles

We saw that A-arguments and P-arguments behave alike for indexing:

both DOI and DSI are **favoured by referential prominence**

It turns out that the generalization is still broader – for example, it applies to recipient arguments, e.g.

(35) Ait Seghrouchen Berber (cited after Souag 2014: 1-2)

- a. *Wši-x aysum i wmušš.*
 give.PFV-1SG meat to cat
 ‘I gave meat to the/a cat’. (Guerssel 1995:115)
- b. *Wši-x =as aysum i wmušš.*
 give.PFV-1SG =3SG.DAT meat to cat
 ‘I gave meat to the/*a cat’. (Guerssel 1995: 115)

It also seems to apply to possessors, e.g.

(36) Khanty (Nikolaeva 1999: 14; 52; cited by Siewierska 2004: 152)

- a. *luw xo:t-al-na*
 he house-3SG-LOC
 ‘in his house’ (personal pronoun possessor)
- b. *Juwan xo:t-na*
 Juwan house-LOC
 ‘in Juwan’s house’ (full nominal possessor)

(37) Welsh (Roberts 1999: 622; cited by Siewierska 2004: 153)

- a. *ei=wraig o*
 3SG=wife he
 ‘his wife’ (personal pronoun possessor)
- b. *(*ei=wraig Gwyn*
 (3SG=)wife Gwyn
 ‘Gwyn’s wife’ (full nominal possessor)

Thus:

differential indexing does not generally behave like differential flagging – it does not show the mirror-image behaviour

7. Concluding remarks

I conclude that differential argument flagging and differential argument indexing must have different explanations

- they show different cross-linguistic patterns

This is not a new conclusion – very similar observations have been made by Iemmolo (2011):

“Overt coding on topical direct objects is indeed **motivated by the unexpectedness of highly topical/topic-worthy lexical direct objects**, because of the low frequency of such NPs in direct object position. By contrast, indexation is **correlated with a high degree of topicality of the referent it indexes**. In other words, overt coding serves to signal that a lexical NP fulfilling the role of direct object is highly topical, while indexation is naturally associated with topical NPs and signals the discourse prominence of such direct objects.” (Iemmolo 2011: 268)

Likewise Witzlack-Makarevich & Seržant (2018: 25-26):

“While agreement or **indexing is “a topic related phenomenon”** as Givón (1976: 185) puts it (cf. also Kibrik 2011), **flagging is not related to topichood or information-structure in general**, but rather to semantic argument roles and various dependency relations between a head and its dependent.”

References

- Aissen, Judith. 1999. Markedness and subject choice in Optimality Theory. *Natural Language & Linguistic Theory* 17. 673–711.
- Aissen, Judith. 2003. Differential object marking: Iconicity vs. economy. *Natural Language & Linguistic Theory* 21(3). 435–483.
- Akter, Mohammed Zahid. 2022. *A comprehensive description of Pangkhua: An endangered Tibeto-Burman language of Bangladesh*. (Thesis.) (<https://ses.library.usyd.edu.au/handle/2123/29852>)
- Anagnostopoulou, Elena. 2006. Clitic doubling. In Everaert, Martin & van Riemsdijk, Henk (eds.), *The Blackwell companion to syntax*, pp. Oxford: Blackwell.
- Berghäll, Liisa. 2015. *A grammar of Mauwake* (Studies in Diversity Linguistics). Berlin: Language Science Press. (<http://langsci-press.org/catalog/book/67>)
- Borsley, Robert D & Tallerman, Maggie & Willis, David. 2007. *The syntax of Welsh*. Cambridge: Cambridge University Press.
- Bossong, Georg. 1985. *Differentielle Objektmarkierung in den neuiranischen Sprachen*. Tübingen: Narr. (<https://zenodo.org/record/4697660>)
- Bossong, Georg. 1991. Differential object marking in Romance and beyond. In Kibbee, Douglas & Wanner, Dieter (eds.), *New analyses in Romance linguistics*, 143–170. Amsterdam: Benjamins. (https://www.rose.uzh.ch/dam/jcr:fffff-c23e-37d9-0000-00006e1a9200/Bossong_80.pdf)
- Chandra, Pritha & Kaur, Gurmeet. 2017. Differential subject marking and person licensing condition. *Concentric: Studies in Linguistics* 43(1). 25–46. (doi:[10.6241/concentric.ling.43.1.02](https://doi.org/10.6241/concentric.ling.43.1.02))
- Compens, Paul. 2022. *Differential object indexing in Bulgarian: The role of discourse prominence and predictability*. Köln: Universität zu Köln. (PhD dissertation.) (<http://www.uni-koeln.de/>)
- Comrie, Bernard. 1979. Definite and animate direct objects: A natural class. *Linguistica Silesiana* 3. 13–21. (doi:[10.5281/zenodo.3862697](https://doi.org/10.5281/zenodo.3862697))
- Comrie, Bernard. 1989. *Language universals and linguistic typology: Syntax and morphology*. Oxford: Blackwell.
- Coon, Jessica & Preminger, Omer. 2017. Split ergativity is not about ergativity. In Coon, Jessica & Massam, Diane & Travis, Lisa Demena (eds.), *The Oxford handbook of ergativity*, 0. Oxford: Oxford University Press. (doi:[10.1093/oxfordhb/9780198739371.013.10](https://doi.org/10.1093/oxfordhb/9780198739371.013.10))
- 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))
- Dalrymple, Mary & Nikolaeva, Irina. 2011. *Objects and information structure*. Cambridge: Cambridge University Press.
- de Hoop, Helen & de Swart, Peter (eds.). 2009. *Differential subject marking* (Studies in Natural Language and Linguistic Theory 72). Dordrecht: Springer.
- Deal, Amy Rose. 2016. Plural exponence in the Nez Perce DP: A DM analysis. *Morphology* 26(3–4). 313–339. (doi:[10.1007/s11525-015-9277-9](https://doi.org/10.1007/s11525-015-9277-9))
- Dixon, R.M.W. 1994. *Ergativity*. Cambridge: Cambridge University Press.
- Fábregas, Antonio. 2013. Differential Object Marking in Spanish: state of the art. *Borealis: An International Journal of Hispanic Linguistics* 2(2). 1–80. (doi:[10.7557/1.2.2.2603](https://doi.org/10.7557/1.2.2.2603))
- Fominyam, Henry Zamchang & Georgi, Doreen. 2021. Subject marking in Awiing. *Studies in African Linguistics* 50(1). 82–115. (doi:[10.32473/sal.v50i1.118528](https://doi.org/10.32473/sal.v50i1.118528))
- Haspelmath, Martin. 2013. Argument indexing: A conceptual framework for the syntax of bound person forms. In Bakker, Dik & Haspelmath, Martin (eds.), *Languages across boundaries: Studies in memory of Anna Siewierska*, 197–226. Berlin: De Gruyter Mouton. (<https://zenodo.org/record/1294059>)
- Haspelmath, Martin. 2015. Descriptive scales versus comparative scales. In Bornkessel-Schlesewsky, Ina & Malchukov, Andrej L. & Richards, Marc (eds.), *Scales and hierarchies: A cross-disciplinary perspective*, 45–58. Berlin: De Gruyter Mouton.
- Haspelmath, Martin. 2019a. Indexing and flagging, and head and dependent marking. *Te Reo* 62(1). 93–115. (doi:[10.17617/2.3168042](https://doi.org/10.17617/2.3168042))
- Haspelmath, Martin. 2019b. Differential place marking and differential object marking. *STUF – Language Typology and Universals* 72(3). 313–334.
- Haspelmath, Martin. 2021a. Explaining grammatical coding asymmetries: Form-frequency correspondences and predictability. *Journal of Linguistics* 57(3). 605–633. (doi:[10.1017/S0022226720000535](https://doi.org/10.1017/S0022226720000535))
- Haspelmath, Martin. 2021b. Role-reference associations and the explanation of argument coding splits. *Linguistics* 59(1). 123–174. (doi:[10.1515/ling-2020-0252](https://doi.org/10.1515/ling-2020-0252))
- Haspelmath, Martin. 2026. Ambiguity avoidance vs. expectation sensitivity as functional factors in grammatical change (to appear). In Gvozdanović, Jadranka (ed.), *(A volume with papers from ICHL Heidelberg)*.
- Iemmolo, Giorgio. 2011. *Towards a typological study of differential object marking and differential object indexation*. Pavia: University of Pavia. (PhD dissertation.) (<https://zenodo.org/record/4388380>)

- Irimia, Monica Alexandrina. 2020. Variation in differential object marking: On some differences between Spanish and Romanian. *Open Linguistics* 6(1). 424–462. (doi:[10.1515/opli-2020-0110](https://doi.org/10.1515/opli-2020-0110))
- Just, Erika. 2022. *A functional approach to differential indexing: Combining perspectives from typology and corpus linguistics*. Leiden: Leiden University. (PhD dissertation.)
- Just, Erika. 2023. Differential indexing and information structure management. In Barotto, Alessandra & Mattioli, Simone (eds.), *Discourse phenomena in typological perspective* (Studies in Language Companion Series), 269–293. Amsterdam: Benjamins.
- Just, Erika. 2024. A structural and functional comparison of differential A and P indexing. *Linguistics* 62(2). 295–321. (doi:[10.1515/ling-2021-0124](https://doi.org/10.1515/ling-2021-0124))
- Kabatek, Johannes & Obrist, Philipp & Wall, Albert. 2021. The third wave of studies on DOM in Romance: An introduction to this volume. In Kabatek, Johannes & Obrist, Philipp & Wall, Albert (eds.), *Differential Object Marking in Romance*, 1–20. Berlin: De Gruyter. (doi:[10.1515/9783110716207-001](https://doi.org/10.1515/9783110716207-001))
- Kalin, Laura. 2018. Licensing and differential object marking: The view from Neo-Aramaic. *Syntax* 21(2). 112–159. (doi:[10.1111/synt.12153](https://doi.org/10.1111/synt.12153))
- Kalin, Laura & Weisser, Philipp. 2019. Asymmetric DOM in coordination: A problem for movement-based approaches. *Linguistic Inquiry* 50(3). 662–676. (doi:https://doi.org/10.1162/ling_a_00298)
- Kayne, Richard S. 2000. *Parameters and universals*. New York: Oxford University Press.
- Koenig, Jean Pierre & Michelson, Karin. 2015. Invariance in argument realization: The case of Iroquoian. *Language* 91. 1–47.
- Lambrecht, Knud. 2000. When subjects behave like objects: An analysis of the merging of S and O in Sentence-Focus Constructions across languages. *Studies in Language. International Journal sponsored by the Foundation "Foundations of Language"* 24(3). 611–682. (doi:[10.1075/sl.24.3.06lam](https://doi.org/10.1075/sl.24.3.06lam))
- Lockwood, Hunter T. & Macaulay, Monica. 2012. Prominence hierarchies. *Language and Linguistics Compass* 6(7). 431–446. (doi:[10.1002/lnc3.345](https://doi.org/10.1002/lnc3.345))
- Malchukov, Andrej. 2008. Animacy and asymmetries in differential case marking. *Lingua* 118. 203–221.
- Merlan, Francesca. 1982. *Mangarayi*. Amsterdam: North-Holland.
- Moravcsik, Edith A. 1978. On the case marking of objects. In Greenberg, Joseph H (ed.), *Universals of human language*, vol. 4: Syntax, 249–289. Stanford: Stanford University Press. (<https://zenodo.org/record/4688136>)
- Morimoto, Yukiko. 2002. Prominence mismatches and Differential Object Marking in Bantu. *Proceedings of the LFG02 Conference*, 292–314. Stanford: CSLI Publications. (<http://csli-publications.stanford.edu/>)
- Nikolaeva, Irina. 1999. Object agreement, grammatical relations, and information structure. *Studies in Language* 23(2). 331–376. (doi:[10.1075/sl.23.2.05nik](https://doi.org/10.1075/sl.23.2.05nik))
- Ouhalla, Jamal. 1993. Subject-extraction, negation and the antiagreement effect. *Natural Language & Linguistic Theory* 11(3). 477–518. (doi:[10.1007/BF00993167](https://doi.org/10.1007/BF00993167))
- Polinsky, Maria & Comrie, Bernard. 1999. Agreement in Tsez. *Folia Linguistica* 33(1–2). 109–130. (doi:[10.1515/flin.1999.33.1-2.109](https://doi.org/10.1515/flin.1999.33.1-2.109))
- Schapper, Antoinette. 2022. *A grammar of Bunaq* (Mouton Grammar Library). Berlin: De Gruyter, Mouton.
- Schikowski, Robert & Iemmolo, Giorgio. 2015. *Commonalities and differences between differential object marking and indexing (working paper)*. Zurich: University of Zurich. (doi:[10.5281/zenodo.7738049](https://doi.org/10.5281/zenodo.7738049))
- Siewierska, Anna. 2004. *Person*. Cambridge: Cambridge University Press.
- Souag, Lameen. 2014. The development of dative agreement in Berber: Beyond nominal hierarchies. *Transactions of the Philological Society*. (doi:[10.1111/1467-968X.12049](https://doi.org/10.1111/1467-968X.12049))
- Stolz, Thomas & Levkovych, Nataliya & Urdze, Aina. 2017. When zero is just enough ... In support of a Special Toponymic Grammar in Maltese. *Folia Linguistica* 51(2). 453–482. (doi:[10.1515/flin-2017-0016](https://doi.org/10.1515/flin-2017-0016))
- Tournadre, Nicolas. 1995. Tibetan ergativity and the trajectory model. In Nishi, Y. & Matisoff, James A & Nagano, Yasuhiko (eds.), *New horizons in Tibeto-Burman morphosyntax*, 261–275. Osaka: National Museum of Ethnology.
- Witzlack-Makarevich, Alena & Seržant, Ilja A. 2018. Differential argument marking: An introduction. In Seržant, Ilja & Witzlack-Makarevich, Alena (eds.), *The diachronic typology of differential argument marking*. Berlin: Language Science Press.
- Zhou, Chenlei. 2022. *Zhoutun* (Routledge World Languages). London: Routledge.

Nonverbal clause constructions:

3A. Overview

MARTIN HASPELMATH

Max Planck Institute for Evolutionary Anthropology

1. Seven initial examples (see Haspelmath 2025a)

- clause types that lack a typical verb, as in (1)-(7).
- in the literature, these are often called “**copula(r) clauses**” (e.g. Declerck 1988; Mikkelsen 2011)
 - but clauses expressing **possession** such as (6) and (7) are also included in the present discussion.
- Perhaps the best-known label for all these clause types is “**nonverbal predication**” (e.g. Hengeveld 1992; Roy 2013; Overall et al. 2018; Creissels et al. 2026).

(1) **classificational construction**

Lee is a baker.

(2) **equational construction**

Kim is my mother.

(3) **attributional construction**

The bird is small.

(4) **predlocative construction** **(predicative-locative)**

The bird is on the roof.

(5) **existential construction**

There is a bird on the roof.

(6) **predpossessive construction** **(predicative-possessive)**

I have a boat.

(7) **appertentive construction** **(‘belong’ construction)**

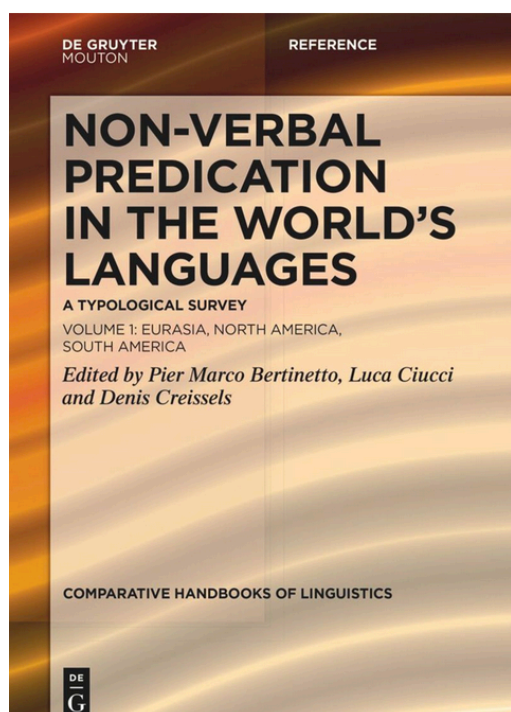
The boat belongs to me.

- Nonverbal clauses have been studied since Clark (1978) and Higgins (1973) (with Lyons 1967 as an important precursor).
- Only copulaless “nominal clauses” were of interest in earlier times, e.g. Meillet 1906) (also called “equational clause”; Sebeok 1943), but an uninterrupted tradition of cross-
- In typological linguistics, key earlier books were Hengeveld (1992), Heine (1997) and Stassen (1997) (see also Freeze 1992 in a generative framework)

2. Nonverbal clauses vs. nonverbal predication

- I use the cover term NONVERBAL CLAUSE CONSTRUCTION for (1)-(7), rather than “nonverbal predication”, because not all of the construction types involve **predication** (see below on “predication”).
- the general concepts and terms used here are comparative concepts (Haspelmath 2018) and not universal categories that are **instantiated** in different languages.
- In terms of the distinction between construction-functions and (construction-)strategies made by Croft (2022) and Haspelmath (2025b), the construction types in (1)-(7) are **construction-functions**, defined independently of their formal properties.

NOTE: Very recently, a new “comparative handbook” has been published, on “non-verbal predication”, edited by Bertinetto, Ciucci and Creissels



Denis Creissels, Pier Marco Bertinetto and Luca Ciucci

1 Non-verbal predication: An analytical framework

Abstract: This chapter outlines, with empirical data drawn from several languages, a theoretical framework accounting for the various manifestations of non-verbal predication. After defining non-verbal predication, it itemizes the different types of non-verbal predication according to: (i) the morphosyntactic nature of the non-verbal predicate (nominal, adjectival, adverbial); (ii) the types of predicative marking (copula construction, juxtaposition, predicative inflection); (iii) the functional domains in which the use of non-verbal predication is particularly prominent (inclusion vs. identity predica-

The introduction text is called by the editors an “*analytical framework*” in the title, and a “*theoretical framework*” in the abstract; and in the concluding section, they say that they offer “a consistent set of *descriptive* concepts”.

But if *description/analysis* is what linguists do when they deal with particular languages ([p-linguistics](#)), then their framework is not “descriptive/analytical”, but **comparative**.

And if a *theory* is an explanatory account from which hypotheses can be derived, their framework is not theoretical, but remains at the **methodological** level, just like my 2025a paper.

3. Duonominal constructions: classificational and equational

duonominal: two nominal expressions that are put in some kind of correspondence

- two well-known subtypes: **classificational** and **equational** clauses
(Creissels et al. : *inclusion* vs. *identity predication*)

Classificational and equational clauses have the same coding properties in most languages, using the same copula and the same argument marking.

(8) Lezgian (Nakh-Dagestanian; Haspelmath 1993: 311)

a. **classificational** (or **inclusional**)

Зи буба кешииш я.
Zi buba kešiš ja.
my father priest COP
'My father is a minister.'

b. **equational** (or **identificational**)

Ви ктаб им я.
Wi ktab im ja.
your book this.one COP
'Your book is this one.'

(9) Welsh (Borsley et al. 2007: 130)

two different copulas!

a. **classificational**

Mae Caerdydd yn ddinas hardd.
be.PRS.3SG Cardiff PRED city beautiful
'Cardiff is a beautiful city.'

b. **equational**

Caerdydd yw prifddinas Cymru.
Cardiff be.PRS.3SG capital Wales
'Cardiff is the capital of Wales.'

(10) Egyptian Arabic (Eid 1991: 41)

a. **classificational**

no copulas

ʕAli mudarris
Ali teacher
'Ali is a teacher.'

b. **equational**

Miḥammad huwwa ʕAli. (**Miḥammad Ø ʕAli.*)
Mohammed COP Ali
'Mohammed is Ali.'

Duonominal clauses are sometimes known as clauses with a “**predicate nominal**” or “nominal predicate” (e.g. Payne 1997: Ch. 6), and this term is quite appropriate for the classificational subtype. Classificational clauses are therefore often called **predicational** (e.g. Higgins 1973; 1979; Mikkelsen 2005).

Equational clauses, too, are often subsumed under “non-verbal predication”, **but it is not clear that they can be said to contain a “predicate”!**

According to Stassen (1997: 12), predication is “the application of a general concept to a particular entity”, and this is not (necessarily) the case in equational clauses (Stassen 1997: 108).

Equational clause constructions fall into four main subtypes:

- (12) a. **characterizational clause**
Kim is the new statistics professor.
- b. **specificational clause**
The new statistics professor is Kim.
- c. **deictic-identificational clause**
That is my brother.
- d. **tautotic clause** (cf. Greek *tautótēs* ‘identity’)
Dr Jekyll is Mr Hyde.

4. Attributional constructions

An ATTRIBUTIONAL CLAUSE CONSTRUCTION expresses the attribution of a property to a subject referent, as already seen in (3) above. There are three more examples in (13)-(15).

- (13) Polish
Marek jest uprzejmy.
 Marek is polite
 ‘Marek is polite.’
- (14) Cantonese (Matthews & Yip 2011: 179)
Léih go jái hóu gōu.
 you CLF son very tall
 ‘Your son is tall.’
- (15) Kolyma Yukaghir (Maslova 2003: 68)
Čumu omo-te-j.
 all good-FUT-3SG
 ‘Everything will be good.’

The property words used in attributional constructions are typically called “adjectives”, but in quite a few languages, **they behave much like typical verbs** and are treated as “stative verbs” (e.g. in Cantonese; see also Francis & Matthews 2005),

BUT: for comparative purposes, it is best to use the terms *noun*, *verb* and *adjective* in a **notional sense** (Stassen 1997: 14; Pustet 2003: 28; Haspelmath 2023), so we may say that all of (13)-(15) contain “**adjectival predicates**”.

5. Predpossessive and appertentive constructions

Four types of clauses having to do with location and possession:

(4) **predlocative construction** (predicative-locative)

The bird is on the roof.

(5) **existential construction**

There is a bird on the roof.

(6) **predpossessive construction** (predicative-possessive)

I have a boat.

(7) **appertentive construction** ('belong' construction)

The boat belongs to me.

PREDPOSSESSIVE ('I have a boat') and APPERTENTIVE ('The boat belongs to me') clause constructions are analogous to existential and predlocative constructions:

(22)

	possessional	locational
definite locatum/possessum	appertentive (<i>The boat is mine.</i>)	predlocative (<i>The boat is on the shore.</i>)
indefinite locatum/possessum	predpossessive (<i>I have a boat.</i>)	existential (<i>There is a boat on the shore.</i>)

Predpossessive constructions are famous for being expressed by fairly different strategies in different languages (e.g. Heine 1997; Creissels 2020).

Transpossessive strategies make use of a transitive verb ('have'), compossessive strategies use a comitative flag ('with') on the possessum, and locpossessive strategies use a locative flag on the possessor ('on, at').

(23) **transpossessive strategy**: Persian

من پول زیادی دارم
Man pul ziyâde dâram.
 I much money have.
 'I have a lot of money.'

(24) **compossessive strategy**: Hausa

Bàlki ta-nàa dà kùjèeruu màasu kyâu.
 Balki 3F.SG-be with chairs having beauty
 'Balki has nice chairs.' (Lit. 'Balki is with nice chairs.') (Abdoulaye 2006: 1122)

(25) **locpossessive strategy**: Finnish

Peka-lla on auto.
 Pekka-ADESS is car.SG.NOM
 'Pekka has a car.' (Lit. 'A car is at Pekka'; Creissels 2013: 468)

There are also two salient strategies for appertentive clauses (see Stolz & Levkovich 2019):

(26) *oblique-possessor appertentive strategy*: Russian

Mašin-a prinadležit Miš-e.
car-NOM belongs Misha-DAT
'The car belongs to Misha.'

(27) *independent-possessor appertentive strategy*: Spanish

Esta canasta es mía.
this basket is mine.
'This basket is mine.' (Stolz & Levkovich 2019: 326)

6. The role of “predicates” in nonverbal clause constructions

- since Hengeveld (1992), the term *nonverbal predication* has become common as a cover term for most of the constructions discussed here.

- In earlier times, these constructions were typically called “*be* constructions” (cf. Verhaar (ed.) 1967-1972) or “copula clauses”, but:

- many languages or constructions lack a copula
- predpossessive clauses are not included

- But why “nonverbal predication”? What is a **predicate** and a **predication**?

In **classificational** (*Lee is a baker*), **attributinal** (*The bord is small*), **predlocative** (*The bird is in the tree*), and **appertentive** constructions (*The boat is mine*), it is clear that there is a subject and a predicate.

Dixon (2010: 163) rejects this usage and prefers to treat adjective, nominal and locative phrase as “copula complements”, i.e. as kinds of arguments, regarding the copula itself as the predicate

But in the other three construction types (equational, existential and predpossessive), it is not clear what the **predicate** might be. Many linguists have of course tried to extend the notions of “subject” and “predicate” to these clauses, too, but no common understanding has emerged.

Freeze (2001: 945) claims that **the locative phrase is the subject** in existential clauses like (28):

(28) Russian

V gorode byl doktor.
in town was doctor
'There was a doctor in town',

But many other linguists would insist that the existent phrase ('the doctor') must be the subject.

→ it is probably best to say that

- some of the nonverbal clause constructions are **predicational** (exhibiting a topic-comment structure),
- while others are **nonpredicational** (as is done by Croft 2022).

predicational	appertentive (‘belong’)		
	predlocative		
	attributional		
	classificational	duonominal	
nonpredicational	equational	characterizational	
		specificational	
		deictic- identificational	
	existential	temporary-location existential	
		permanent-presence existential	
	predpossessive (‘have’)		
	hyparctic (‘exist’)		

Table 1: The main types of nonverbal clause constructions and subtypes

equational		
classificational	ascriptive	
attributional		
predlocative	locational	locopossessional
existential		
predpossessive	possessional	
appertentive		
hyparctic		

Table 2: Nonverbal clause construction types and super-types

7. Copulas

Many nonverbal clause types have copulas:

(28) Gyeli (Bantu)

Àdà àà ngèlèné.

Ada COP.G1 teacher(G1)

'Ada is a teacher.' (Grimm 2021: 440)

(29) Jamaican (English-based creole)

Fi-dem moni de pan di tiebl.

of-them money LCOP on the table

'Their money is on the table.' (Farquharson 2013)

A copula is best defined as **a form that indicates a stative link between the two argument positions of an equational, ascriptive or locational clause** (i.e. a predlocative or existential clause).

Comments:

(1) maybe it would be more natural to say that a copula is *any* linking element in a nonverbal clause?

But ‘have’ verbs in predpossessionive clauses and ‘belong’ verbs in an appertentive clauses (e.g. 26) are never called copulas – copulas must be restricted to equational, ascriptive and locational clauses, even though these are not a natural class of clause types.

(2) a copula is defined as **a kind of form**, and since a form is overt and segmental by definition, this definition **excludes the possibility of a “zero copula”**.

This also means that tonal marking of nonverbal predication, as has been reported for a number of African languages by Gibson et al. (2019: §2.3), does not count as “copula”.

(30) Herero (Bantu; Gibson et al. 2019: §2.3)

a. *ò-tjì-havérò*
AUG-G7-chair
‘chair’

b. *ó-tjì-havérò*
AUG-G7-chair
‘it is a chair’

(3) a form can be **affixal**, so **affixal copulas** such as -y in (31) from Turkish are included in the definition as well (Creissels et al. (2026) call such forms “copulative affixes”).

(31) Turkish

Satıcı-y-dı-m.
seller-COP-PST-1SG
‘I was a seller.’ (Kornfilt 1997: 57)

But affixal markers that only indicate person-number, tense and so on **do not count as copulas**; an example is -t in (32). In Bertinetto et al. (2019), such constructions are called **predicative inflection**.

(32) Erzya (Uralic)

Ist’amo šumbra ćora-t di a šim-at.
such wealthy man-2SG and NEG drink-2SG
‘You are such a wealthy man and you don’t drink.’ (Turunen 2011: 148)

Such forms may look quite verb-like in these languages!

(4) the definition does not say anything about a copula’s “**verbal**” or “**nonverbal**” **nature**. Many copulas are quite verb-like, but others clearly look non-verb-like (e.g. Egyptian Arabic *huwwa* in (10b), which derives from a personal pronoun ‘he’).

(Creissels et al. 2026 say that copulas are special because they do not assign semantic roles to the subject, so they do not count as “verbal” in this sense.)

8. Juxtaposition tendencies

In Haspelmath (2025a), I formulate no generalizations, but Creissels et al. (2026: 19) note a number of cross-linguistic tendencies for the use of juxtaposition (as opposed to a copula):

Across the languages that adopt this strategy, the juxtaposition construction is more frequent:

- in nominal predication than in adverbial predication;
- with third-person arguments than with arguments representing speech act participants;
- in clauses expressing a TAM value of the type commonly labeled “indicative/realis present” than in clauses expressing other TAM values;
- in plain-locational predication than in inverse-locational (or in “existential”) predication (see Section 4.4);
- with nominal predicates than with adjectival predicates (Dryer 2007: 229–230);
- in independent clauses than in subordinate clauses;
- in clauses that do not involve overt marking of information structure than in clauses in which a term is overtly focalized.

9. Equational clauses: Characterizational and specificational subtypes

Equational clauses have two subtypes: CHARACTERIZATIONAL and SPECIFICATIONAL clauses. The difference between them is best characterized by means of the notions of TOPIC and COMMENT, and of VARIABLE and VALUE (cf. Huddleston & Pullum 2002: 53; Mikkelsen 2011: 1810)

(35) a. **characterizational clause**

Warsaw_{VALUE} is the capital of Poland_{VARIABLE}.

b. **specificational clause**

The capital of Poland_{VARIABLE} is Warsaw_{VALUE}.

Characterizational and specificational clauses do not simply **equate** two referents (“Dr. Jekyll is Mr. Hyde”), but **combine a value with a variable**.

In a **characterizational clause**, the **value nominal is the topic** and the **variable is the comment**, so it is appropriate in contexts such as (36a-b).

- (36) a. A: *I heard about our new colleague Kim. Do you know more about her?*
 B: *Kim is the new statistics professor.* (cf. (12a) above)
- b. A: *I heard about a city called Warsaw. Can you tell me more about it?*
 B: *Warsaw is the capital of Poland.* (= (35a); Stassen 1997: 103)

The VALUE NOMINAL can be a name, but the VARIABLE NOMINAL must be a nominal headed by a common noun which serves to *characterize* the referent in a non-exhaustive way.

In a **specificational clause**, the **variable nominal is the topic** and the **value is the comment**: The clause serves to uniquely SPECIFY the value of a variable.

Thus, specificational clauses are appropriate in contexts such as (37a-b), where the variable (the role of statistics professor, the role of capital) is topical.

- (37) a. A: *Do you know how the vacancy in statistics was filled?*
 B: *The new statistics professor is Kim.* (cf. (12b) above)
- b. A: *I know about Poland, but can you remind me of its capital?*
 B: *The capital of Poland is Warsaw.* (= (35b); Stassen 1997: 103)

In English, the topic is usually the precopular nominal and the comment is in postcopular position, so the literature often equates *characterizational* with value-initial copula clauses, and *specificational* with value-final copula clauses. However, the following dialogue is possible in English as well, with (38/B) basically equivalent to (37b/B).

- (38) A: *What is the capital of Poland?*
 B: *WARSAW^{VALUE} is the capital of Poland^{VARIABLE}.*

Here the value comes first, but the context makes it clear that the postposed variable is topical, so we are dealing with a specificational clause.

Many clauses can be interpreted in two ways:

- (39) *My best friend is the new statistics professor.*

If this occurs in the context “Can you tell me about your best friend, Kim?”, it is characterizational, but if it occurs in the context of “I wonder who is your best friend”, it is specificational.

NOTE: the “grammatical subject” properties are not uniform in specificational clauses across languages!

In English, the precopular variable nominal controls number agreement of the copula verb, but in Portuguese, **agreement is controlled by the postcopular value nominal**,

- (40) *O problema são os teus pais.*
 the problem.SG COP.PL the your parents.PL
 ‘The problem is your parents.’ (Heycock 2012: 211)

Portuguese does not allow singular *é* ‘is’ here, suggesting that the postcopular nominal is the “subject”, but the English counterpart does not allow plural *are* (**The problem are your parents*).

Similarly, Italian shows person agreement controlled by the postnominal value nominal, while French is like English in not allowing this.

- (41) a. Italian (Moro 1997: 219)
Il colpevole sono io.
 the culprit am I
 ‘The culprit is me.’
 b. French (Heycock 2012: 211-213)
Le coupable c’ est moi. (**Le coupable suis moi.*)
 the culprit that is me the culprit am I
 ‘The culprit is me.’

German is like Portuguese and Italian in this regard, not like English or French (Heycock 2012).

There are a number of “copula-like” contexts where a reversal of value and variable is not possible:

- (42) a. *I regard Kim as my best friend. / *I regard my best friend as Kim.*
 b. *I consider Kim (to be) my best friend. / I consider my best friend to be Kim.*

The variable nominal is treated as **an inanimate entity** with respect to resumptive pronouns, for example in left dislocation constructions and tag question constructions in English (Mikkelsen 2005: Ch. 5)

- (43) a. *The tallest girl in the class, that is Mariamu.* (**... she is Mariamu*)
 b. *The tallest girl in the class, she is Kenyan.* (**... that is Kenyan*)
 (44) a. *Your best friend is Kim, isn’t it?* (**... isn’t she?*)
 b. *Your best friend is a statistician, isn’t she?* (**... isn’t it?*)

10. Deictic-identificational clauses

Equational clauses in quite a few languages show special properties when one of the nominals is a demonstrative, e.g.

(47) German

Das ist meine Schwester.
 that.N is my.F sister(F)
 ‘That is my sister.’

And in English, the demonstrative *that* can be used without *one* only in such deictic-identificational clauses (cf. *That (one) is my bike.* vs. *I bought that **one** last year.*)

In quite a few languages, there are special **predicative-identifying demonstratives** that are only used in such contexts (Diessel 1997; 1999: §4.3; Killian 2021: §5).

(48) Pohnpeian (Oceanic; Rehg & Sohl 1981: 143, 150)

a. *Met pahn mengila.*
 DEM FUT wither
 ‘This will wither.’

b. *Iet noumw naipen.*
 PRED.DEM your knife
 ‘This/here is your knife.’

Such predicative demonstratives are similar to OSTENSIVE MARKERS such as French *voici* and *voilà* and Hebrew *hine*, illustrated in (49b).

(49) a. French

Voilà Omar.
 OST Omar
 ‘There is Omar/There comes Omar.’

b. Hebrew

Hine Natan.
 OST Nathan
 ‘Here is Nathan.’ (Izre’el 2023: 121)

11. Concluding remark

It should always be kept in mind that the categories of particular languages do not always coincide with salient comparative concepts.

For example, the **English There-is Construction** does not coincide with existential constructions, because it can also be used in hyparctic clauses (e.g. *There is beer without alcohol*).

And the Polish verb *mieć* ‘have’ can be used both in predpossession constructions and in negative existential constructions (*nie ma* ‘there isn’t’), which is not a natural class.

The term *nonverbal clause construction* itself is here defined as the set of clause types given in Table 1 and Table 2, which are all defined in functional-semantic terms. It is a set of construction-functions, not strategies.

predicational	appertentive ('belong')		
	predlocative		
	attributional		
	classificational		duonominal
nonpredicational	equational	characterizational	
		specificational	
		deictic- identificational	
	existential	temporary-location existential	
		permanent-presence existential	
	predpossessive ('have')		
hyparctic ('exist')			

Table 1: The main types of nonverbal clause constructions and subtypes

equational		
classificational	ascriptive	
attributional		
predlocative	locational	locopossessional
existential	possessional	
predpossessive		
appertentive		
hynarcctic		

Table 2: Nonverbal clause construction types and super-types

References

- Abdoulaye, Mahamane L. 2006. Existential and possessive predications in Hausa. *De Gruyter Mouton* 44(6). 1121–1164. (doi:10.1515/LING.2006.037)
- Bentley, Delia. 2017. Copular and existential constructions. In Dufter, Andreas & Stark, Elisabeth (eds.), *Manual of Romance morphosyntax and syntax*, 332–366. Berlin: De Gruyter. (doi 10.1515/9783110377088-009) (Accessed September 13, 2021.)
- Bentley, Delia & Ciconte, Francesco Maria & Cruschina, Silvio. 2015. *Existentials and locatives in Romance dialects of Italy*. Oxford: Oxford University Press.
- Benveniste, Émile. 1960. «Etre» et «avoir» dans leurs fonctions linguistiques. *Bulletin de la Société de Linguistique de Paris* 55. (Also in: *Problèmes de linguistique générale* 1, 187–207. Paris: Gallimard, 1966)
- Bertinetto, Pier Marco & Ciucci, Luca & Farina, Margherita. 2019. Two types of morphologically expressed non-verbal predication. *Studies in Language*. John Benjamins 43(1). 120–194. (doi:10.1075/sl.17013.ber)
- Bickerton, Derek. 1981. *Roots of language*. Ann Arbor: Karoma Publishers [Republished in 2016 by Language Science Press: <https://langsci-press.org/catalog/book/91>].
- Borsley, Robert D & Tallerman, Maggie & Willis, David. 2007. *The syntax of Welsh*. Cambridge: Cambridge University Press.
- Chappell, Hilary & Lü, Shanshan. 2022. A semantic typology of location, existence, possession and copular verbs: Areal patterns of polysemy in Mainland East and Southeast Asia. *Linguistics* 60(1). 1–82. (doi:10.1515/ling-2021-0219)
- Clark, Eve V. 1978. Locational: Existential, locative, and possessive constructions. In Greenberg, Joseph H. (ed.), *Universals of human language, Vol. 4: Syntax*, 85–126. Stanford: Stanford University Press.
- Creissels, Denis. 2013. Control and the evolution of possessive and existential constructions. In van Gelderen, Elly & Cennamo, Michela (eds.), *Argument Structure in Flux: the Naples-Capri papers*, 461–476. Amsterdam: Benjamins.
- Creissels, Denis. 2019. Inverse-locational predication in typological perspective. *Italian Journal of Linguistics* 31(2). 37–106.
- Creissels, Denis. 2020. La prédication existentielle et les constructions transpossessives. In Bottineau, Tatiana (ed.), *La prédication existentielle dans les langues naturelles : Valeurs et repérages, structures et modalités*, 17–41. Paris: Presses de l'Inalco.
- Creissels, Denis & Bertinetto, Pier Marco & Ciucci, Luca. 2025. Non-verbal predication in the world's languages: An overview. In: *Oxford handbook of non-verbal predication* (to appear).
- Croft, William. 2022. *Morphosyntax: Constructions of the world's languages*. Cambridge: Cambridge University Press. (<http://www.unm.edu/~wcroft/WACpubs.html>)

- Curnow, Timothy Jowan. 2000. Towards a cross-linguistic typology of copula constructions. In Henderson, John (ed.), *Proceedings of the 1999 Conference of the Australian Linguistic Society*, 1–9. (<http://www.als.asn.au>)
- Declerck, Renaat. 1983. 'It is Mr. Y' or 'he is Mr. Y'? *Lingua* 59(2). 209–246. (doi:10.1016/0024-3841(83)90064-5)
- Declerck, Renaat. 1988. *Studies on copular sentences, clefts and pseudo-clefts*. Leuven: Leuven University Press & Foris.
- den Dikken, Marcel & O'Neill, Teresa. 2017. Copular constructions in syntax. *Oxford Research Encyclopedia of Linguistics*. Oxford: Oxford University Press. (doi:10.1093/acrefore/9780199384655.013.137)
- Diessel, Holger. 1997. Predicative demonstratives. *Annual Meeting of the Berkeley Linguistics Society* 23(1). 72–82.
- Diessel, Holger. 1999. *Demonstratives: Form, function, and grammaticalization*. Amsterdam: Benjamins.
- Dik, Simon C. 1997. *The theory of Functional Grammar: 2 volumes*. Berlin: Mouton de Gruyter.
- Dixon, R. M. W. 2010. *Basic linguistic theory: Volume 2*. Oxford: Oxford University Press.
- Dryer, Matthew S. 2007. Clause types. In Shopen, Timothy (ed.), *Language typology and syntactic description, Vol. 1*, 224–275. Second Edition. Cambridge: Cambridge University Press.
- Eid, Mushira. 1991. Verbless sentences in Arabic and Hebrew. In Comrie, Bernard & Eid, Mushira (eds.), *Perspectives on Arabic Linguistics III: Papers from the Third Annual Symposium on Arabic Linguistics*, 31–61. Amsterdam: Benjamins.
- Falk, Yehuda N. 2006. *Subjects and Universal Grammar: An explanatory theory*. Cambridge: Cambridge University Press.
- Farquharson, Joseph T. 2013. Jamaican structure dataset. In Michaelis, Susanne Maria & Maurer, Philippe & Haspelmath, Martin & Huber, Magnus (eds.), *Atlas of Pidgin and Creole Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. (<https://apics-online.info/contributions/8>)
- Foley, William A. & Van Valin, Robert D. 1977. On the viability of the notion of "subject" in universal grammar. *Berkeley Linguistics Society* 3. 293–320.
- Forker, Diana. 2013. *A grammar of Hiniuk*. Berlin: De Gruyter Mouton.
- Francis, Elaine J. & Matthews, Stephen. 2005. A multi-dimensional approach to the category 'verb' in Cantonese. *Journal of Linguistics* 41(2). 269–305. (doi:10.1017/S0022226705003270)
- Freeze, Ray. 2001. Existential constructions. In Haspelmath, Martin & König, Ekkehard & Oesterreicher, Wulf & Raible, Wolfgang (eds.), *Language typology and language universals: An international handbook*, 941–953. Berlin: Mouton de Gruyter.
- Geist, Ljudmila. 2007. Predication and equation in copular sentences: Russian vs. English. In Comorovski, Ileana & von Heusinger, Klaus (eds.), *Existence: Semantics and syntax*, 79–105. Dordrecht: Springer.
- Gibson, Hannah & Guérois, Rozenn & Marten, Lutz. 2019. Variation in Bantu copula constructions. In Arche, María J. & Fábregas, Antonio & Marín, Rafael (eds.), *The grammar of copulas across languages*, 213–242. Oxford: Oxford University Press.
- Grimm, Nadine. 2021. *A grammar of Gyeli* (Comprehensive Grammar Library 2). Berlin: Language Science Press.
- Hamari, Arja. 2022. The negative existential cycle in Moksha Mordvin: From a negative existential into a negative auxiliary. In Veselinova, Ljuba & Hamari, Arja (eds.), *The Negative Existential Cycle*, 325–356. Berlin: Language Science Press. (DOI: 10.5281/zenodo.7353613)
- Haspelmath, Martin. 1993. *A grammar of Lezgian* (Mouton Grammar Library 9). Berlin: Mouton de Gruyter.
- Haspelmath, Martin. 2011. On S, A, P, T, and R as comparative concepts for alignment typology. *Linguistic Typology* 15(3). 535–567.
- 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)
- Haspelmath, Martin. 2018. How comparative concepts and descriptive linguistic categories are different. In Van Olmen, Daniël & Mortelmans, Tanja & Brisard, Frank (eds.), *Aspects of linguistic variation: Studies in honor of Johan van der Auwera*, 83–113. Berlin: De Gruyter Mouton. (<https://zenodo.org/record/3519206>)
- Haspelmath, Martin. 2020a. The structural uniqueness of languages and the value of comparison for description. *Asian Languages and Linguistics* 1(2). 346–366.
- Haspelmath, Martin. 2020b. The morph as a minimal linguistic form. *Morphology* 30(2). 117–134. (doi:10.1007/s11525-020-09355-5)
- Haspelmath, Martin. 2023. Word class universals and language-particular analysis. In van Lier, Eva (ed.), *The Oxford handbook of word classes*, 15–40. Oxford: Oxford University Press.
- Haspelmath, Martin. 2024. Construction-functions and construction-strategies. To appear.
- Haspelmath, Martin & Buchholz, Oda. 1998. Equative and similative constructions in the language of Europe. In Auwera, van der (ed.), *Adverbial constructions in the languages of Europe* (Empirical Approaches to Language Typology/EUROTYP, 20-3), 277–334. Berlin.
- Hedberg, Nancy & Potter, David. 2010. Equative and predicational copulas in Thai. *Berkeley Linguistics Society* 36(1). 144–157. (doi:10.3765/bls.v36i1.3908)
- Heine, Bernd. 1997. *Possession: Cognitive sources, forces, and grammaticalization*. Cambridge: Cambridge University Press.
- Hengeveld, Kees. 1992. *Non-verbal predication: Theory, typology, diachrony*. Berlin: Mouton de Gruyter.

- Heycock, Caroline. 2012. Specification, equation, and agreement in copular sentences. *Canadian Journal of Linguistics/Revue canadienne de linguistique*. Cambridge University Press 57(2). 209–240. (doi:10.1017/S0008413100004758)
- Heycock, Caroline & Kroch, Anthony. 1999. Pseudocleft connectedness: Implications for the LF Interface Level. *Linguistic Inquiry* 30(3). 365–397. (doi:10.1162/002438999554110)
- Higgins, F. R. 1973. *The pseudo-cleft construction in English*. Cambridge, MA: MIT dissertation. (<https://dspace.mit.edu/handle/1721.1/12988>)
- Higgins, F. R. 1979. *The pseudo-cleft construction in English*. New York: Garland.
- Huddleston, Rodney & Pullum, Geoffrey K. (eds.). 2002. *The Cambridge grammar of the English language*. Cambridge: Cambridge University Press.
- Izre'el, Shlomo. 2023. The syntax of presentative constructions: The view from colloquial Israeli Hebrew. In T.I. Davidjuk, I.I. Isaev, Ju.V. Mazurova, S. G. Tatevosov & O.I. Fedorova (eds.), *Jazyk kak on est': Sbornik statej k 60-letiju Andreja Aleksandroviča Kibrika*, 121–128. Moscow: Buki Vedi.
- Jespersen, Otto. 1924. *The philosophy of grammar*. London: Allen & Unwin.
- Killian, Don. 2021. Towards a typology of predicative demonstratives. *Linguistic Typology*. (doi:10.1515/lingty-2021-2078)
- Kim, Jong-Bok. 2016. Copular constructions and asymmetries in the specificational pseudocleft constructions in Korean. *Language and Linguistics*. SAGE Publications Ltd 17(1). 89–112. (doi:10.1177/1606822X15586672)
- Koch, Peter. 2012. Location, existence, and possession: A constructional-typological exploration. *Linguistics* 50(3). 533–603. (doi:10.1515/ling-2012-0018)
- Kornfilt, Jaklin. 1997. *Turkish*. London: Routledge.
- Kuno, Susumu & Wongkhamthong, Preya. 1981. Characterizational and Identificational Sentences in Thai. *Studies in Language* 5(1). 65–109. (doi:10.1075/sl.5.1.04kun)
- Lefebvre, Claire & Brousseau, Anne-Marie. 2002. *A grammar of Fongbe*. Berlin: Mouton de Gruyter.
- Locker, Ernst. 1954. Etre et avoir: Leurs expressions dans les langues. *Anthropos* 49(3/4). 481–510.
- Lyons, John. 1967. A note on possessive, existential and locative sentences. *Foundations of Language* 3(4). 390–396.
- Lyons, John. 1968. *Introduction to theoretical linguistics*. Cambridge: Cambridge University Press.
- Lyons, John. 1977. *Semantics*. Cambridge: Cambridge University Press.
- Maslova, Elena. 2003. *A grammar of Kolya Yukaghir* (Mouton Grammar Library 27). Berlin: Mouton de Gruyter.
- Matthews, Stephen & Yip, Virginia. 2011. *Cantonese: A comprehensive grammar*. London: Routledge.
- McGregor, William B. 2001. Non-verbal predicative possession in Nyulnyulan languages. In Simpson, Jane & Nash, David & Laughren, Mary & Austin, Peter & Alpher, Barry (eds.), *Forty years on: Ken Hale and Australian languages*, 337–352. Canberra: Australian National University. (Accessed September 13, 2021.)
- McNally, Louise. 2016. Existential sentences crosslinguistically: Variations in form and meaning. *Annual Review of Linguistics* 2(1). 211–231. (doi:10.1146/annurev-linguistics-011415-040837)
- Meillet, Antoine. 1906. La phrase nominale en indo-européen. *Mémoires de la Société de Linguistique de Paris* 14. 1–26.
- Michaelis, Susanne Maria & Rosalie, Marcel. 2013. Seychelles Creole structure dataset. In Michaelis, Susanne Maria & Maurer, Philippe & Haspelmath, Martin & Huber, Magnus (eds.), *Atlas of Pidgin and Creole Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. (<https://apics-online.info/contributions/56>)
- Mikkelsen, Line. 2005. *Copular clauses: Specification, predication and equation*. Amsterdam: Benjamins.
- Mikkelsen, Line. 2007. On so-called truncated clefts. In Geist, Ljudmila & Rothstein, Björn (eds.), *Kopulaverben und Kopulasätze: Intersprachliche und intrasprachliche Aspekte*, 47–68. Tübingen: Niemeyer.
- Mikkelsen, Line. 2011. Copular clauses. In Maienborn, Claudia & von Heusinger, Klaus & Portner, Paul (eds.), *Semantics: An international handbook of natural language meaning, volume 2*, 1805–1829. Berlin: De Gruyter Mouton.
- Milsark, Gary L. 1977. Toward an explanation of certain peculiarities of the existential construction in English. *Linguistic Analysis* 3(2). 1–29.
- Moro, Andrea. 1997. *The raising of predicates: Predicative noun phrases and the theory of clause structure*. Cambridge: Cambridge University Press.
- Moro, Andrea. 2006. Existential sentences and expletive *there*. In Everaert, Martin & van Riemsdijk, Henk (eds.), *The Blackwell companion to syntax, volume II*, 210–236. Oxford: Blackwell.
- Myler, Neil. 2016. *Building and interpreting possession sentences*. Cambridge MA: MIT Press.
- Nordlinger, Rachel. 1998. *A grammar of Wambaya, Northern Australia*. Canberra: Pacific Linguistics. (<http://dx.doi.org/10.15144/PL-C140>)
- Nordlinger, Rachel & Sadler, Louisa. 2007. Verbless clauses: Revealing the structure within. In Zaenen, Annie E & Simpson, Jane & King, Tracy Holloway & Grimshaw, Jane & Maling, Joan & Manning, Chris (eds.), *Architectures, rules, and preferences: Variations on themes by Joan W. Bresnan*, 139–160. Stanford: CSLI Publications.

- Overall, Simon E. 2018. Nonverbal predicates and copula constructions in Aguaruna (Chicham). In Overall, Simon E. & Vallejos, Rosa & Gildea, Spike (eds.), *Nonverbal predication in Amazonian languages*, 135–161. Amsterdam: Benjamins.
- Overall, Simon E. & Vallejos, Rosa & Gildea, Spike (eds.). 2018. *Nonverbal predication in Amazonian languages*. Amsterdam: Benjamins.
- Padučeva, Elena V. 1985. *Vyskazyvanie i ego sootnesennost' s dejstvitel'nost'ju: Referencial'nye aspekty semantiki mestoimenij*. Moskva: Nauka.
- Paducheva, Elena V. 2008. Locative and existential meaning of Russian быть. *Russian Linguistics* 32(3). 147–158.
- Partee, Barbara H. 2010. Specificational Copular Sentences in Russian and English. *Oslo Studies in Language* 2(1). (doi:10.5617/osla.127) (<https://journals.uio.no/osla/article/view/127>) (Accessed July 19, 2021.)
- Partee, Barbara H. & Borschev, Vladimir. 2007. Existential sentences, BE, and the genitive of negation in Russian. In Comorovski, Ileana & von Heusinger, Klaus (eds.), *Existence: Semantics and syntax*, 147–190. Dordrecht: Springer.
- Payne, Thomas. 1997. *Describing morphosyntax: A guide for field linguists*. Cambridge: Cambridge University Press.
- Petrollino, Sara. 2019. Existential predication in Hamar. *Nordic Journal of African Studies* 28(4). 27–27.
- Pustet, Regina. 2003. *Copulas: Universals in the categorization of the lexicon*. Oxford: Oxford University Press.
- Rehg, Kenneth L. & Sohl, Damian G. 1981. *Ponapean reference grammar*. Honolulu: University of Hawaii Press.
- Rennison, John R. 1997. *Koromfe* (Descriptive Grammar Series). London: Routledge.
- Rothstein, Susan. 1995. Small clauses and copular constructions. In Cardinaletti, Anna (ed.), *Small clauses*, 27–48. New York: Academic Press. (doi:10.1163/9780585492209_003)
- Roy, Isabelle. 2013. *Nonverbal predication: Copular sentences at the syntax-semantics interface*. Oxford: Oxford University Press.
- Sabbagh, Joseph. 2009. Existential sentences in Tagalog. *Natural Language & Linguistic Theory* 27(4). 675–719. (doi:10.1007/s11049-009-9083-3)
- Sebeok, Thomas A. 1943. The equational sentence in Hungarian. *Language* 19(4). 320–327. (doi:10.2307/409935)
- Stassen, Leon. 1997. *Intransitive predication*. Oxford: Oxford University Press.
- Stassen, Leon. 2005. Predicative possession. In Haspelmath, Martin & Dryer, Matthew S. & Gil, David & Comrie, Bernard (eds.), *The world atlas of language structure*, 474–477. Oxford: Oxford University Press. (<http://wals.info/chapter/117>)
- Stassen, Leon. 2009. *Predicative possession*. Oxford: Oxford University Press.
- Stolz, Thomas. 2001. To be with X is to have X: Comitatives, instrumentals, locative, and predicative possession. *Linguistics* 39(2). 321–350. (doi:10.1515/ling.2001.013)
- Stolz, Thomas & Levkovych, Nataliya. 2019. On belonging: Preliminary thoughts on the typology of belong-constructions. In Johanson, Lars & Mazzitelli, Lidia Federica & Nevskaya, Irina (eds.), *Possession in languages of Europe and North and Central Asia*, 313–363. Amsterdam: Benjamins. (<https://benjamins.com/catalog/slcs.206>)
- Tham, Shiao Wei. 2013. Possession as non-verbal predication. *Annual Meeting of the Berkeley Linguistics Society*, vol. 39, 302–316.
- Turunen, Rigina. 2010. *Nonverbal predication in Erzya: Studies on morphosyntactic variation and part of speech distinctions*. Helsinki: University of Helsinki. (Ph.D. dissertation.) (<https://helda.helsinki.fi/handle/10138/19652>) (Accessed September 13, 2021.)
- Vallejos, Rosa. 2016. *A grammar of Kukama-Kukamiria: A language from the Amazon*. Leiden: Brill.
- van Gijn, Erik. 2006. *A grammar of Yurakaré*. Nijmegen: Radboud University. (Ph.D. dissertation.) (http://webdoc.ubn.ru.nl/mono/g/gijn_e_van/gramofyu.pdf)
- Van Valin, Robert D., Jr. 2005. *Exploring the syntax-semantics interface*. Cambridge: Cambridge University Press.
- Verhaar, John W. M. (ed.). 1967-1972. *The verb “be” and its synonyms: Philosophical and grammatical studies (5 volumes, 1967-1972)*. Dordrecht: Reidel.
- Veselinova, Ljuba. 2013. Negative existentials: A cross-linguistic study. *Rivista di Linguistica* 25(1). 107–145.
- Veselinova, Ljuba & Hamari, Arja (eds.). 2022. *The Negative Existential Cycle* (Research on Comparative Grammar). Language Science Press (Research on Comparative Grammar). Berlin: Language Science Press. (Accessed September 13, 2021.)
- Ye, Jingting. 2020. Independent and dependent possessive person forms: Three universals. *Studies in Language* 44(2). 363–406. (doi:10.1075/sl.19020.ye)
- Ye, Jingting. 2021. *Property words and adjective subclasses in the world's languages*. Leipzig: Leipzig University. (Ph D dissertation.)

Nonverbal clause constructions:

3B. Existential constructions

MARTIN HASPELMATH

Max Planck Institute for Evolutionary Anthropology

1. Starting out from a stereotype (asking ChatGPT)

Give a typical existential clause from English, Spanish and Chinese.

Great question — let's look at what an existential clause is first, and then examples from English, Spanish, and Chinese.

◆ What is an existential clause?

An existential clause expresses the existence or presence of something — it answers “there is / there are” in English. It typically introduces something new into the discourse.

🚩 1. English

Example: *There is a book on the table.*

Structure: there + be + noun phrase (+ location / complement)

Explanation: “there” is a dummy subject — it doesn't mean a place, just marks existence.

🇪🇸 2. Spanish

Example: *Hay un libro en la mesa.*

Structure: hay (from haber) + noun phrase (+ location)

Explanation: “hay” literally means “there is / there are”.
Spanish doesn't need a dummy subject; the verb *haber* alone expresses existence.

🇨🇳 3. Chinese (Mandarin)

Example: 桌子上有一本书。
Zhuōzi shàng yǒu yī běn shū.
“There is a book on the table.”

Structure: [Location] + 有 (yǒu ‘have/there is’) + noun phrase

Explanation: The verb 有 functions as an existential verb meaning “there is/are.”
The location phrase comes first.

This is all fine, except the definition: “expresses the existence or presence of something”, which is far too vague.

How can we improve on this? How can we study existential clauses systematically across languages?

2. The main strategies for existential clauses

Existential clauses use a range of fairly different strategies that we may want to have specific terms for, e.g.

(50) **existive-copula strategy**: Tagalog

May mga tao sa labas.

EXV PL person LOC outside

‘There are people outside.’ (Sabbagh 2009: 678)

(51) **transpossessionive-existential strategy**: Seychelles Creole

Ler i annan koudvan zot pa reste lo sa zil.

when 3SG have hurricane they not stay on the island

‘When there is a hurricane they don’t stay on the island.’ (Michaelis & Rosalie 2013)

(52) **prolocative strategy**: Italian

Ci sono molte montagne in Svizzera.

PROLOC are many mountains in Switzerland

‘There are many mountains in Switzerland.’

French has a mixture of the latter two types, making use both of *avoir* ‘have’ and of a prolocative element *y* ‘there’ (e.g. *il y a beaucoup de montagnes en Suisse* ‘there are many mountains in Switzerland’).

Spanish could be said to use the same mixture in the present tense (*ha-y*), but simply *haber* (transpossessionive-existential) in other tenses. However, *haber* no longer means ‘have’, so synchronically, Spanish *haber* is best treated as an existive copula.

The last type is merely characterized by the postposed ordering of the existent, as in (53).

(53) **existent-postposing strategy**: Finnish

Kato-lla on lintu.

roof-ADESS is bird

‘There is a bird on the roof.’ (cf. *Lintu on katolla* ‘The bird is on the roof.’)

And another strategy: no difference between existential and predicative-locative clauses:

(57) Mandinka (Mande)

Wulôo bé yîròo kótò.

dog.DET LCOP tree.DET under

‘The dog is under the tree.’ OR: ‘There is a dog under the tree.’

(Creissels 2019: 51)

3. A possible universal regularity: The existent prefers an object-like position

VO languages: – post-copula position (*there are mountains...*)
 – post-“verbal” position (*il y a des montagnes...*)

SOV languages: – immediately preverbal position

Japanese *Yane-no ue-ni tori-ga imasu.*
 roof-GEN on-LOC bird-NOM be
 ‘There is a bird on the roof.’

Tori-wa yane-no ue-ni imasu.
 bird-TOP roof-GEN on-LOC be
 ‘The bird is on the roof.’

Freeze (1992: 555) gives the following examples from Hindi:

(3) Hindi:

- a. *māī hindustaan-mēē thaa.*
 I India-in COP.Sg.MASC.PAST (Theme Locative V)
 ‘I was in India.’
- b. *kamree-mēē aadmī hai.*
 room-in man COP.3sg.MASC.PRES (Locative Theme V)
 ‘There is a man in the room.’

Existential constructions are semantically quite similar to predicative-locative constructions: there is a **locative phrase** (e.g. ‘on the roof’), there is located element (the **locatum**), and the clause expresses the locational relationship.

(Both can be said to be subtypes of “**locational clauses**”.)

The locatum can be called the **existent phrase** in existential constructions (Haspelmath 2025; also called “pivot” in the literature).

Existential constructions have not been studied very thoroughly from a worldwide perspective (but see Clark 1978; Freeze 1992; Creissels 2019).

But one may propose the following universal:

(1) **The existent-position universal**

If a language prefers to place the existent phrase in a position different from the locatum phrase in a predicative-locative construction, it occurs in a position that is analogous to the position of the object in a transitive clause.

The object-like position is understandable because the existent phrase is indefinite and is normally new information, just as objects typically are.

4. Existential clauses do not express “existence”

Creissels (2019) objects to the term “existential clause”, because the clause type in question cannot be characterized by the notion of “existence”.

(ChatGPT: “An existential clause expresses the existence or presence of something” – but this is not a good definition.)

“if ‘exist(ence)’ is taken with its usual meaning ‘be(ing) an element of the world’, it should be clear that *There is a cat in the tree* does not express existence, since in languages having a verb such as English *exist* or French *exister*, it is impossible to use it to express the same meaning, and conversely, in many languages, clauses expressing pure existence without any reference being made to location, such as *God exists* / *Dieu existe*, cannot be paraphrased by means of the inverse-locational predicator.” (Creissels 2019: 37)

Creissels is right – clauses expressing “pure existence” should not be called “existential clauses”. I have proposed a new term:

hyparctic clauses (based on Greek *hýparxis* (ὑπαρξις) ‘existence’)

A few examples:

(64) Latin

Cogito, ergo sum.

I.think therefore I.exist

‘I think, therefore I am.’ (René Descartes, 1637)

(65) French

Il est des situations dans lesquelles personne n’aimerait se trouver.

it is INDF situations in which nobody would.like self to.find

‘There are situations in which nobody would like to be.’ (Creissels 2019: 44)

(66) Hamar (South Omotic)

Bajó dāa.

fate EXV

‘Fate exists.’ (Petrollino 2019: 8)

(67) Classical Arabic

لَا إِلَهَ إِلَّا اللَّهُ

Laa ʔilaah-a ʔillaa ʔllaahu.

not god-ACCunless Allah

‘There is no deity but God.’ (Qur’an 37:35)

- there are diverse strategies for hyparctic clauses
- there is a lot of overlap with existential clauses, but there are also differences (for example, French *être* (seen in (65)) can be used in hyparctic, but not in existential clauses)
- hyparctic clauses have not been studied systematically yet, but it would be a worthwhile topic for typology

Creissels et al. (2026: 33):

“The importance of a terminological distinction between these two clause types is also acknowledged by Haspelmath (2025), but the solution he proposes is different ... the use of two etymologically equivalent labels (such as “existential” and “hyparctic”) for two distinct constructions does not contribute to more clearly reflecting their differences.”

However, I do not think that etymology matters much in grammatical terminology.

5. Definitions: Locational and possessional clauses

From the beginning (Lyons 1967; Clark 1978), it has been noted that existential clauses are related to predlocative clauses and/or predpossessive clauses in many languages.

(4) **predlocative construction** (predicative-locative)

The bird is on the roof.

(5) **existential construction**

There is a bird on the roof.

(6) **predpossessive construction** (predicative-possessive)

I have a boat.

(7) **appertentive construction**

The boat belongs to me.

(22)

	possessional	locational
definite locatum/possessum	appertentive (<i>The boat is mine.</i>)	predlocative (<i>The boat is on the shore.</i>)
indefinite locatum/possessum	predpossessive (<i>I have a boat.</i>)	existential (<i>There is a boat on the shore.</i>)

• Predlocative vs. existential constructions

A PREDLOCATIVE CONSTRUCTION is a clause construction in which a **definite** subject argument is said to be **located in a place expressed by a locative phrase**:

(16) Wambaya (Australia; Nordlinger 1998: 177)

Janji inyaga jalyu-ni!
dog.NOM that.NOM bed-LOC
‘The dog is on the bed!’

(17) Koromfe (Gur, West Africa; Rennison 1997: 65)

Də wě dāānɛ.
he LCOP at.home
‘He is at home.’

An EXISTENTIAL CONSTRUCTION is a clause in which **an indefinite nominal phrase** (the EXISTENT) is said to be in some location that is **generally expressed by a locative phrase**.

- (18) Logudorese Sardinian (Bentley et al. 2015: 7)

In custu istradone nch' at una creža.

in this road there have.3SG a church

‘In this road there is a church.’ (Lit. ‘It there **has** a church in this road.’)

- (19) Aguaruna (Chicham; Overall 2018: 154)

Utunchat a-yi nunka=num.

problem exist-PST.3SG.DECL land=LOC

‘There was a problem in the land.’

- (20) Wambaya (Australia; Nordlinger 1998: 177)

Garnguji julaji-rdarra gayangga darranggu-ni.

many.NOM bird-GROUP.NOM high tree-LOC

‘There are lots of birds up in the trees.’

The existent is often called **pivot** (following Milsark 1977), and Creissels (2019) calls it the **figure** (contrasting with the *ground* expressed by the locative phrase).

One interesting difference:

– the locative phrase must be overt in a predlocative clause, but can be left implicit in an existential clause:

The coffee is on the table.

**The coffee is.*

There is coffee on the table.

There is coffee.

• temporary location vs. permanent presence

In some languages, there is a clear difference between temporary location and permanent presence of the indefinite nominal, e.g.

- (21) German

a. temporary location of existent

Auf dem Tisch sind Blumen.

on the table are flowers

‘There are flowers on the table.’ (Cf. ?**Auf dem Tisch gibt es Blumen.*)

b. permanent presence of existent

In Thailand gibt es Tiger.

in Thailand gives it tigers

‘There are tigers in Thailand.’ (Cf. ?**In Thailand sind Tiger.*)

Some authors have treated this as an important distinction (e.g. Koch 2012), and Creissels (2019) restricts his attention to **temporary-location existentials**. However, many languages do not treat the two types differently.

6. Copulas and existive forms

Some languages have a linking form that is restricted to existential (or predpossessionive) clauses, such as Spanish *hay* in (33) or Tagalog *may*.

I have proposed the term EXISTIVE for such special forms (called “specialized inverse-locational predicators” by Creissels 2019: 80).

- (33) Spanish
***Hay** mucha gente afuera.*
 EXV many people outside
 ‘There are many people outside.’

Particularly common are negative existives, e.g. Russian *net* ‘there isn’t’, Turkish *yok* ‘there isn’t’ (Veselinova 2013; Veselinova & Hamari 2022).

Note that an existive is a type of copula: it can be defined as a copula that does not occur outside of existential and predpossessionive clauses.

7. “Inverse-locational” and other existential clauses

In Haspelmath (2025a), an EXISTENTIAL CLAUSE is a clause in which an indefinite nominal phrase (the locatum) is said to be in some location.

This definition of the term *existential clause* corresponds to widespread usage (since Lyons 1967; 1968: §8.4; Clark 1978). A few more examples of existential clauses are given in (54)-(56).

- (54) Kukama-Kukamiria (Tupí-Guaraní)
***Emete** ikiratsen tukini=kuara.*
 exist kid hammock=INESS
 ‘There is a kid in the hammock.’ (Vallejos 2016: 345)
- (55) Hinuq (Nakh-Dagestanian)
*Ardel elu-de [?]aši šayt’an-be **zoq’e-s**=ñen.*
 formerly we-LOC many devil-PL be-PST=QUOT
 ‘Formerly in our village there were many devils.’ (Forker 2013: 533)
- (56) English
There was a bird on the roof.
On the roof was a bird.
A BIRD was on the roof.

Some authors have used the term *existential* in a narrower sense, to refer to a subclass of **existential clauses that are “noncanonical” in some way**, “whether due to some aspect of their syntax or the presence of a distinguished lexical item (e.g. Spanish *hay*)” (McNally 2016: 212).

But this is not a homogeneous class, because “what is canonical differs from language to language” (as McNally notes), so it is better to use the traditional term *existential* in a broad functional-semantic sense.

Creissels (2019) examines one subtype of existential clause, which he calls ***inverse-locational predication***.

In this type of clause, there is an episodic spatial relationships involving two concrete entities: a **figure** that has the ability to move, and a **ground** occupying a fixed position in space, and the clause is “**perspectivized**” from **ground to figure**.

There is a cat in the tree.

Inverse-locational clauses contrast with *plain-locational clauses*, where the perspectivization is from figure to ground (e.g.

The cat is in the tree.

Creissels (2019: 41-42) limits the term *inverse-locational* to constructions which are “**specialized**” in the sense that they cannot be derived from predlocative constructions by some generally applicable device such as word order or marking for definiteness or topicality.

In other words, Creissels’ term *inverse-locational* refers to **a specific strategy** (in Croft’s 2022 and Haspelmath’s 2025b sense) to express the locational relationship in question.

Thus, Creissels would not classify an English sentence such as *A BIRD is on the roof* as inverse-locational, and he notes that in some languages, there are sentences which can have both interpretations, e.g.

(57) Mandinka (Mande)

Wùlôo bé yíróo kótò.

dog.DET LCOP tree.DET under

‘The dog is under the tree.’ OR: ‘There is a dog under the tree.’

(Creissels 2019: 51)

A large number of languages (probably more than half of the world’s languages) lack a (dedicated) inverse-locational construction.

Creissels does not require the locatum in inverse-locational clauses to be indefinite, citing French sentences like

Tiens, il y a Jean!

‘Hey, there’s Jean!’

Creissels (2019: §2.4) rightly criticizes Koch (2012) for equating the contrast between predlocative and existential clauses with the contrast between a **thematic/topical locatum** and a **rhematic/focal locatum**.

We can have predlocative sentences with a focal locatum (*THE DOG is on the bed*), and (59)-(60) are existential sentences whose existent is topicalized.

(59) Russian

[Я искал кефир.] Кефира в магазине не было.

[Ja iskal kefir.] Kefir-a v magazine ne byl-o.

I.NOM sought kefir.ACC kefir-GEN in store NEG was-N.SG

‘[I was looking for kefir.] There wasn’t any kefir in the store.’ (Partee & Borschev 2007: 155)

(60) French

Des femmes, il y en avait plusieurs à la réunion.

INDF women it there of.them had several at the meeting

(lit.) ‘Women, there were several of them at the meeting.’ (Creissels 2019: 49)

References

- Abdoulaye, Mahamane L. 2006. Existential and possessive predications in Hausa. *De Gruyter Mouton* 44(6). 1121–1164. (doi:10.1515/LING.2006.037)
- Bentley, Delia. 2017. Copular and existential constructions. In Dufter, Andreas & Stark, Elisabeth (eds.), *Manual of Romance morphosyntax and syntax*, 332–366. Berlin: De Gruyter. (doi 10.1515/9783110377088-009) (Accessed September 13, 2021.)
- Bentley, Delia & Ciconte, Francesco Maria & Cruschina, Silvio. 2015. *Existentials and locatives in Romance dialects of Italy*. Oxford: Oxford University Press.
- Benveniste, Émile. 1960. «Etre» et «avoir» dans leurs fonctions linguistiques. *Bulletin de la Société de Linguistique de Paris* 55. (Also in: *Problèmes de linguistique générale* 1, 187–207. Paris: Gallimard, 1966)
- Bertinetto, Pier Marco & Ciucci, Luca & Farina, Margherita. 2019. Two types of morphologically expressed non-verbal predication. *Studies in Language*. John Benjamins 43(1). 120–194. (doi:10.1075/sl.17013.ber)
- Bickerton, Derek. 1981. *Roots of language*. Ann Arbor: Karoma Publishers [Republished in 2016 by Language Science Press: <https://langsci-press.org/catalog/book/91>].
- Borsley, Robert D & Tallerman, Maggie & Willis, David. 2007. *The syntax of Welsh*. Cambridge: Cambridge University Press.
- Chappell, Hilary & Lü, Shanshan. 2022. A semantic typology of location, existence, possession and copular verbs: Areal patterns of polysemy in Mainland East and Southeast Asia. *Linguistics* 60(1). 1–82. (doi:10.1515/ling-2021-0219)
- Clark, Eve V. 1978. Locational: Existential, locative, and possessive constructions. In Greenberg, Joseph H. (ed.), *Universals of human language, Vol. 4: Syntax*, 85–126. Stanford: Stanford University Press.
- Creissels, Denis. 2013. Control and the evolution of possessive and existential constructions. In van Gelderen, Elly & Cennamo, Michela (eds.), *Argument Structure in Flux: the Naples-Capri papers*, 461–476. Amsterdam: Benjamins.
- Creissels, Denis. 2019. Inverse-locational predication in typological perspective. *Italian Journal of Linguistics* 31(2). 37–106.
- Creissels, Denis. 2020. La prédication existentielle et les constructions transpossessives. In Bottineau, Tatiana (ed.), *La prédication existentielle dans les langues naturelles : Valeurs et repérages, structures et modalités*, 17–41. Paris: Presses de l’Inalco.
- Creissels, Denis & Bertinetto, Pier Marco & Ciucci, Luca. 2025. Non-verbal predication in the world’s languages: An overview. In: *Oxford handbook of non-verbal predication* (to appear).
- Croft, William. 2022. *Morphosyntax: Constructions of the world’s languages*. Cambridge: Cambridge University Press. (<http://www.unm.edu/~wcroft/WACpubs.html>)
- Curnow, Timothy Jowan. 2000. Towards a cross-linguistic typology of copula constructions. In Henderson, John (ed.), *Proceedings of the 1999 Conference of the Australian Linguistic Society*, 1–9. (<http://www.als.asn.au>)
- Declerck, Renaat. 1983. ‘It is Mr. Y’ or ‘he is Mr. Y’? *Lingua* 59(2). 209–246. (doi:10.1016/0024-3841(83)90064-5)
- Declerck, Renaat. 1988. *Studies on copular sentences, clefts and pseudo-clefts*. Leuven: Leuven University Press & Foris.

- den Dikken, Marcel & O'Neill, Teresa. 2017. Copular constructions in syntax. *Oxford Research Encyclopedia of Linguistics*. Oxford: Oxford University Press. (doi:10.1093/acrefore/9780199384655.013.137)
- Diessel, Holger. 1997. Predicative demonstratives. *Annual Meeting of the Berkeley Linguistics Society* 23(1). 72–82.
- Diessel, Holger. 1999. *Demonstratives: Form, function, and grammaticalization*. Amsterdam: Benjamins.
- Dik, Simon C. 1997. *The theory of Functional Grammar: 2 volumes*. Berlin: Mouton de Gruyter.
- Dixon, R. M. W. 2010. *Basic linguistic theory: Volume 2*. Oxford: Oxford University Press.
- Dryer, Matthew S. 2007. Clause types. In Shopen, Timothy (ed.), *Language typology and syntactic description, Vol. 1*, 224–275. Second Edition. Cambridge: Cambridge University Press.
- Eid, Mushira. 1991. Verbless sentences in Arabic and Hebrew. In Comrie, Bernard & Eid, Mushira (eds.), *Perspectives on Arabic Linguistics III: Papers from the Third Annual Symposium on Arabic Linguistics*, 31–61. Amsterdam: Benjamins.
- Falk, Yehuda N. 2006. *Subjects and Universal Grammar: An explanatory theory*. Cambridge: Cambridge University Press.
- Farquharson, Joseph T. 2013. Jamaican structure dataset. In Michaelis, Susanne Maria & Maurer, Philippe & Haspelmath, Martin & Huber, Magnus (eds.), *Atlas of Pidgin and Creole Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. (<https://apics-online.info/contributions/8>)
- Foley, William A. & Van Valin, Robert D. 1977. On the viability of the notion of “subject” in universal grammar. *Berkeley Linguistics Society* 3. 293–320.
- Forker, Diana. 2013. *A grammar of Hinuq*. Berlin: De Gruyter Mouton.
- Francis, Elaine J. & Matthews, Stephen. 2005. A multi-dimensional approach to the category ‘verb’ in Cantonese. *Journal of Linguistics* 41(2). 269–305. (doi:10.1017/S0022226705003270)
- Freeze, Ray. 2001. Existential constructions. In Haspelmath, Martin & König, Ekkehard & Oesterreicher, Wulf & Raible, Wolfgang (eds.), *Language typology and language universals: An international handbook*, 941–953. Berlin: Mouton de Gruyter.
- Geist, Ljudmila. 2007. Predication and equation in copular sentences: Russian vs. English. In Comorovski, Ileana & von Heusinger, Klaus (eds.), *Existence: Semantics and syntax*, 79–105. Dordrecht: Springer.
- Gibson, Hannah & Guérois, Rozenn & Marten, Lutz. 2019. Variation in Bantu copula constructions. In Arche, Maria J. & Fábregas, Antonio & Marín, Rafael (eds.), *The grammar of copulas across languages*, 213–242. Oxford: Oxford University Press.
- Grimm, Nadine. 2021. *A grammar of Gyeli* (Comprehensive Grammar Library 2). Berlin: Language Science Press.
- Hamari, Arja. 2022. The negative existential cycle in Moksha Mordvin: From a negative existential into a negative auxiliary. In Veselinova, Ljuba & Hamari, Arja (eds.), *The Negative Existential Cycle*, 325–356. Berlin: Language Science Press. (DOI: 10.5281/zenodo.7353613)
- Haspelmath, Martin. 1993. *A grammar of Lezgian* (Mouton Grammar Library 9). Berlin: Mouton de Gruyter.
- Haspelmath, Martin. 2011. On S, A, P, T, and R as comparative concepts for alignment typology. *Linguistic Typology* 15(3). 535–567.
- 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)
- Haspelmath, Martin. 2018. How comparative concepts and descriptive linguistic categories are different. In Van Olmen, Daniël & Mortelmans, Tanja & Brisard, Frank (eds.), *Aspects of linguistic variation: Studies in honor of Johan van der Auwera*, 83–113. Berlin: De Gruyter Mouton. (<https://zenodo.org/record/3519206>)
- Haspelmath, Martin. 2020a. The structural uniqueness of languages and the value of comparison for description. *Asian Languages and Linguistics* 1(2). 346–366.
- Haspelmath, Martin. 2020b. The morph as a minimal linguistic form. *Morphology* 30(2). 117–134. (doi:10.1007/s11525-020-09355-5)
- Haspelmath, Martin. 2023. Word class universals and language-particular analysis. In van Lier, Eva (ed.), *The Oxford handbook of word classes*, 15–40. Oxford: Oxford University Press.
- Haspelmath, Martin. 2024. Construction-functions and construction-strategies. To appear.
- Haspelmath, Martin & Buchholz, Oda. 1998. Equative and similative constructions in the language of Europe. In Auwera, van der (ed.), *Adverbial constructions in the languages of Europe* (Empirical Approaches to Language Typology/EUROTYP, 20-3), 277–334. Berlin.
- Hedberg, Nancy & Potter, David. 2010. Equative and predication copulas in Thai. *Berkeley Linguistics Society* 36(1). 144–157. (doi:10.3765/bls.v36i1.3908)
- Heine, Bernd. 1997. *Possession: Cognitive sources, forces, and grammaticalization*. Cambridge: Cambridge University Press.
- Hengeveld, Kees. 1992. *Non-verbal predication: Theory, typology, diachrony*. Berlin: Mouton de Gruyter.
- Heycock, Caroline. 2012. Specification, equation, and agreement in copular sentences. *Canadian Journal of Linguistics/Revue canadienne de linguistique*. Cambridge University Press 57(2). 209–240. (doi:10.1017/S0008413100004758)
- Heycock, Caroline & Kroch, Anthony. 1999. Pseudocleft connectedness: Implications for the LF Interface Level. *Linguistic Inquiry* 30(3). 365–397. (doi:10.1162/002438999554110)

- Higgins, F. R. 1973. *The pseudo-cleft construction in English*. Cambridge, MA: MIT dissertation. (<https://dspace.mit.edu/handle/1721.1/12988>)
- Higgins, F. R. 1979. *The pseudo-cleft construction in English*. New York: Garland.
- Huddleston, Rodney & Pullum, Geoffrey K. (eds.). 2002. *The Cambridge grammar of the English language*. Cambridge: Cambridge University Press.
- Izre'el, Shlomo. 2023. The syntax of presentative constructions: The view from colloquial Israeli Hebrew. In T.I. Davidjuk, I.I. Isaev, Ju.V. Mazurova, S. G. Tatevosov & O.I. Fedorova (eds.), *Jazyk kak on est': Sbornik statej k 60-letiju Andreja Aleksandroviča Kibrika*, 121–128. Moscow: Buki Vedi.
- Jespersen, Otto. 1924. *The philosophy of grammar*. London: Allen & Unwin.
- Killian, Don. 2021. Towards a typology of predicative demonstratives. *Linguistic Typology*. (doi:10.1515/lingty-2021-2078)
- Kim, Jong-Bok. 2016. Copular constructions and asymmetries in the specificational pseudocleft constructions in Korean. *Language and Linguistics*. SAGE Publications Ltd 17(1). 89–112. (doi:10.1177/1606822X15586672)
- Koch, Peter. 2012. Location, existence, and possession: A constructional-typological exploration. *Linguistics* 50(3). 533–603. (doi:10.1515/ling-2012-0018)
- Kornfilt, Jaklin. 1997. *Turkish*. London: Routledge.
- Kuno, Susumu & Wongkhamthong, Preya. 1981. Characterizational and Identificational Sentences in Thai. *Studies in Language* 5(1). 65–109. (doi:10.1075/sl.5.1.04kun)
- Lefebvre, Claire & Brousseau, Anne-Marie. 2002. *A grammar of Fongbe*. Berlin: Mouton de Gruyter.
- Locker, Ernst. 1954. Etre et avoir: Leurs expressions dans les langues. *Anthropos* 49(3/4). 481–510.
- Lyons, John. 1967. A note on possessive, existential and locative sentences. *Foundations of Language* 3(4). 390–396.
- Lyons, John. 1968. *Introduction to theoretical linguistics*. Cambridge: Cambridge University Press.
- Lyons, John. 1977. *Semantics*. Cambridge: Cambridge University Press.
- Maslova, Elena. 2003. *A grammar of Kolyma Yukaghir* (Mouton Grammar Library 27). Berlin: Mouton de Gruyter.
- Matthews, Stephen & Yip, Virginia. 2011. *Cantonese: A comprehensive grammar*. London: Routledge.
- McGregor, William B. 2001. Non-verbal predicative possession in Nyulnyulan languages. In Simpson, Jane & Nash, David & Laughren, Mary & Austin, Peter & Alpher, Barry (eds.), *Forty years on: Ken Hale and Australian languages*, 337–352. Canberra: Australian National University. (Accessed September 13, 2021.)
- McNally, Louise. 2016. Existential sentences crosslinguistically: Variations in form and meaning. *Annual Review of Linguistics* 2(1). 211–231. (doi:10.1146/annurev-linguistics-011415-040837)
- Meillet, Antoine. 1906. La phrase nominale en indo-européen. *Mémoires de la Société de Linguistique de Paris* 14. 1–26.
- Michaelis, Susanne Maria & Rosalie, Marcel. 2013. Seychelles Creole structure dataset. In Michaelis, Susanne Maria & Maurer, Philippe & Haspelmath, Martin & Huber, Magnus (eds.), *Atlas of Pidgin and Creole Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. (<https://apics-online.info/contributions/56>)
- Mikkelsen, Line. 2005. *Copular clauses: Specification, predication and equation*. Amsterdam: Benjamins.
- Mikkelsen, Line. 2007. On so-called truncated clefts. In Geist, Ljudmila & Rothstein, Björn (eds.), *Kopulaverben und Kopulasätze: Intersprachliche und intrasprachliche Aspekte*, 47–68. Tübingen: Niemeyer.
- Mikkelsen, Line. 2011. Copular clauses. In Maienborn, Claudia & von Stechow, Klaus & Portner, Paul (eds.), *Semantics: An international handbook of natural language meaning, volume 2, 1805–1829*. Berlin: De Gruyter Mouton.
- Milsark, Gary L. 1977. Toward an explanation of certain peculiarities of the existential construction in English. *Linguistic Analysis* 3(2). 1–29.
- Moro, Andrea. 1997. *The raising of predicates: Predicative noun phrases and the theory of clause structure*. Cambridge: Cambridge University Press.
- Moro, Andrea. 2006. Existential sentences and expletive *there*. In Everaert, Martin & van Riemsdijk, Henk (eds.), *The Blackwell companion to syntax, volume II*, 210–236. Oxford: Blackwell.
- Myler, Neil. 2016. *Building and interpreting possession sentences*. Cambridge MA: MIT Press.
- Nordlinger, Rachel. 1998. *A grammar of Wambaya, Northern Australia*. Canberra: Pacific Linguistics. (<http://dx.doi.org/10.15144/PL-C140>)
- Nordlinger, Rachel & Sadler, Louisa. 2007. Verbless clauses: Revealing the structure within. In Zaenen, Annie E & Simpson, Jane & King, Tracy Holloway & Grimshaw, Jane & Maling, Joan & Manning, Chris (eds.), *Architectures, rules, and preferences: Variations on themes by Joan W. Bresnan*, 139–160. Stanford: CSLI Publications.
- Overall, Simon E. 2018. Nonverbal predicates and copula constructions in Aguaruna (Chicham). In Overall, Simon E. & Vallejos, Rosa & Gildea, Spike (eds.), *Nonverbal predication in Amazonian languages*, 135–161. Amsterdam: Benjamins.
- Overall, Simon E. & Vallejos, Rosa & Gildea, Spike (eds.). 2018. *Nonverbal predication in Amazonian languages*. Amsterdam: Benjamins.
- Padučeva, Elena V. 1985. *Vyskazyvanie i ego sootnesennost' s dejstvitel'nost'ju: Referencial'nye aspekty semantiki mestoimenij*. Moskva: Nauka.
- Paducheva, Elena V. 2008. Locative and existential meaning of Russian быть. *Russian Linguistics* 32(3). 147–158.

- Partee, Barbara H. 2010. Specificational Copular Sentences in Russian and English. *Oslo Studies in Language* 2(1). (doi:10.5617/osla.127) (<https://journals.uio.no/osla/article/view/127>) (Accessed July 19, 2021.)
- Partee, Barbara H. & Borschev, Vladimir. 2007. Existential sentences, BE, and the genitive of negation in Russian. In Comorovski, Ileana & von Heusinger, Klaus (eds.), *Existence: Semantics and syntax*, 147–190. Dordrecht: Springer.
- Payne, Thomas. 1997. *Describing morphosyntax: A guide for field linguists*. Cambridge: Cambridge University Press.
- Petrollino, Sara. 2019. Existential predication in Hamar. *Nordic Journal of African Studies* 28(4). 27–27.
- Pustet, Regina. 2003. *Copulas: Universals in the categorization of the lexicon*. Oxford: Oxford University Press.
- Rehg, Kenneth L. & Sohl, Damian G. 1981. *Ponapean reference grammar*. Honolulu: University of Hawaii Press.
- Rennison, John R. 1997. *Koromfe* (Descriptive Grammar Series). London: Routledge.
- Rothstein, Susan. 1995. Small clauses and copular constructions. In Cardinaletti, Anna (ed.), *Small clauses*, 27–48. New York: Academic Press. (doi:10.1163/9780585492209_003)
- Roy, Isabelle. 2013. *Nonverbal predication: Copular sentences at the syntax-semantics interface*. Oxford: Oxford University Press.
- Sabbagh, Joseph. 2009. Existential sentences in Tagalog. *Natural Language & Linguistic Theory* 27(4). 675–719. (doi:10.1007/s11049-009-9083-3)
- Sebeok, Thomas A. 1943. The equational sentence in Hungarian. *Language* 19(4). 320–327. (doi:10.2307/409935)
- Stassen, Leon. 1997. *Intransitive predication*. Oxford: Oxford University Press.
- Stassen, Leon. 2005. Predicative possession. In Haspelmath, Martin & Dryer, Matthew S. & Gil, David & Comrie, Bernard (eds.), *The world atlas of language structure*, 474–477. Oxford: Oxford University Press. (<http://wals.info/chapter/117>)
- Stassen, Leon. 2009. *Predicative possession*. Oxford: Oxford University Press.
- Stolz, Thomas. 2001. To be with X is to have X: Comitatives, instrumentals, locative, and predicative possession. *Linguistics* 39(2). 321–350. (doi:10.1515/ling.2001.013)
- Stolz, Thomas & Levkovych, Nataliya. 2019. On belonging: Preliminary thoughts on the typology of belong-constructions. In Johanson, Lars & Mazzitelli, Lidia Federica & Nevskaya, Irina (eds.), *Possession in languages of Europe and North and Central Asia*, 313–363. Amsterdam: Benjamins. (<https://benjamins.com/catalog/slcs.206>)
- Tham, Shiao Wei. 2013. Possession as non-verbal predication. *Annual Meeting of the Berkeley Linguistics Society*, vol. 39, 302–316.
- Turunen, Rigina. 2010. *Nonverbal predication in Erzya: Studies on morphosyntactic variation and part of speech distinctions*. Helsinki: University of Helsinki. (Ph.D. dissertation.) (<https://helda.helsinki.fi/handle/10138/19652>) (Accessed September 13, 2021.)
- Vallejos, Rosa. 2016. *A grammar of Kukama-Kukamiria: A language from the Amazon*. Leiden: Brill.
- van Gijn, Erik. 2006. *A grammar of Yurakaré*. Nijmegen: Radboud University. (Ph.D. dissertation.) (http://webdoc.ubn.ru.nl/mono/g/gijn_e_van/gramofyu.pdf)
- Van Valin, Robert D., Jr. 2005. *Exploring the syntax-semantics interface*. Cambridge: Cambridge University Press.
- Verhaar, John W. M. (ed.). 1967-1972. *The verb “be” and its synonyms: Philosophical and grammatical studies (5 volumes, 1967-1972)*. Dordrecht: Reidel.
- Veselinova, Ljuba. 2013. Negative existentials: A cross-linguistic study. *Rivista di Linguistica* 25(1). 107–145.
- Veselinova, Ljuba & Hamari, Arja (eds.). 2022. *The Negative Existential Cycle* (Research on Comparative Grammar). *Language Science Press* (Research on Comparative Grammar). Berlin: Language Science Press. (Accessed September 13, 2021.)
- Ye, Jingting. 2020. Independent and dependent possessive person forms: Three universals. *Studies in Language* 44(2). 363–406. (doi:10.1075/sl.19020.ye)
- Ye, Jingting. 2021. *Property words and adjective subclasses in the world’s languages*. Leipzig: Leipzig University. (Ph D dissertation.)

Nonverbal clause constructions:

3C. Adjectives vs. verbs vs. nouns

MARTIN HASPELMATH

Max Planck Institute for Evolutionary Anthropology

1. Attributional clauses with “adjectival verbs”: A discussion of Creissels et al. (2026)

I said earlier that “an **attributional clause construction** expresses the attribution of a **property** to a subject referent:

(13) Polish

Marek jest uprzejmy.

Marek is polite

‘Marek is polite.’

(14) Cantonese (Matthews & Yip 2011: 179)

Léih go jái hóu gōu.

you CLF son very tall

‘Your son is tall.’

(15) Kolyma Yukaghir (Maslova 2003: 68)

Čumu omo-te-j.

all good-FUT-3SG

‘Everything will be good.’

But is a sentence such as (15) “non-verbal”?

Dryer (2007: 227) does not think so (see also Overall et al. 2018: 3):

“Adjectival predicates in English are nonverbal because English treats adjectives as a distinct word class from verbs. In many languages, however, the words expressing meanings associated with adjectives in English **are simply verbs**. In such languages, adjectival predicates are thus **not a kind of nonverbal predicate**, but simply a type of intransitive verbal predicate.”

e.g. in Cree:

- | | | | |
|------|---|--|---|
| (10) | a. nipā-w
sleep-3SG
‘he/she sleeps’ | b. nipā-wak
sleep-3PL
‘they sleep’ | c. ni-nipā-n
1-sleep-1SG
‘I sleep’ |
| (11) | a. mišikiti-w
big-3SG
‘he/she is big’ | b. mišikiti-wak
big-3PL
‘they are big’ | c. ni-mišikiti-n
1-big-1SG
‘I am big’ |

But what exactly is an “adjective”?

Recall that terms such as adjective have been used in two ways: as **construction-functions** (= property concept words) and as **construction-strategies**. Dryer (2007) and Overall (2018) evidently use “adjective” for a kind of strategy. But how is this strategy defined in general terms?

“Adjectives” are widely known among typologists to be **“less universal than nouns and verbs”** (e.g. Panagiotidis 2015: 41; see also e.g. Evans & Levinson 2009: 435), and it has been a perennial question how they should be treated in language description and in typology.

How should we deal with predicative property-concept words as in (1)-(3), which lack a copula?

(1) Russian

Этот цветок красный.
Ètot cvetok krasnyj.
 ‘This flower is red.’

(2) Japanese

この花は赤い。
Kono hana wa akai.
 ‘This flower is red.’

(3) Haitian Creole

Flè sa a wouj.
 ‘This flower is red.’

Are the words for ‘red’ in these languages “adjectives”?

At least Japanese *akai* and Haitian Creole *wouj* are **verb-like** also in that they can **combine with a past-tense marker**

Japanese	<i>aka-kutta</i> red-PST ‘was red’
Haitian	<i>té wouj</i> PST red ‘was red’

Are they verbs, or are they adjectives after all?

Creissels et al. (2026: 10) define adjectives as **roots that can occur adnominally “without additional devices”**:

“Our proposal is that the notion of adjective is relevant for the languages having a class of underived lexemes characterized by the possibility of fulfilling the function of adnominal modifier **without the additional devices** that are necessary for underived lexemes belonging to other morphosyntactic classes to fulfill the same function. Such lexemes can be designated as primary adjectives.”

→ Russian *krasnyj*, Japanese *akai* and Haitian *wouj* are **adjectives**

(can be used adnominally without relativizers:
‘red flower’ is *krasnyj cvetok*, *akai hana*, and *flè wouj*)

BUT:

– This definition fails to exclude **nouns which can be used adnominally**, and this is not so uncommon.

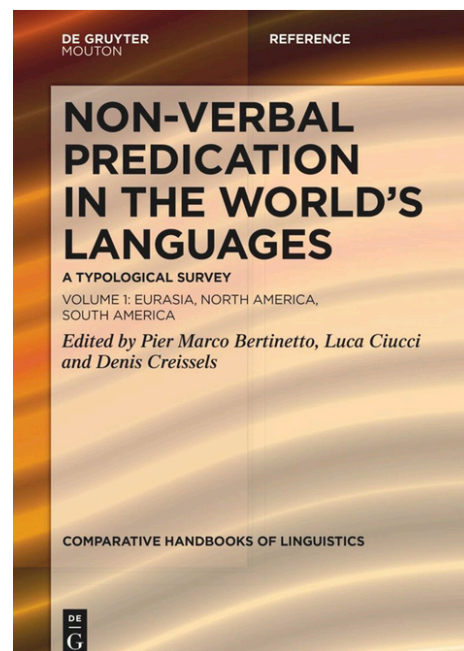
In a sample of 72 languages worldwide that have noun juxtaposition of some kind, Mizuno (2024) found 24 languages that use **juxtaposition for adnominal possession** (e.g. Nivkh *ətək χaj* [father pigeon] ‘father’s pigeon’)

– in these languages, nouns can “fulfill the function of adnominal modifier without additional devices”

Moreover, some languages do not even need special “devices” for relative clauses; for example, Japanese has

tobu dōbutsu
fly animal
‘animal that flies’

without a relativizer. Does this make *tobu* ‘fly’ an “adjective” in Japanese?



(In the definition, there is a strange formulation “without the additional devices that are necessary for underived lexemes belonging to **other** morphosyntactic classes” – but which are the “other classes”? It seems that we can know this only if we already know what an adjective is.)

Creissels et al. refer to the most recent handbook article about adjectives, Beck (2023), who defines “adjective” as follows:

“The term ‘adjective’ will be used to describe a lexical–syntactic class of word that contains primarily expressions of property concepts and whose syntactic distribution can be distinguished from that of other word classes in that it shows **a specialization for adnominal modification**.” (Beck 2023: 366)

Here the term “specialization” may refer to the absence of a “special adnominal device” (i.e. an absence of a relativizer, as in Russian, Japanese and Haitian Creole), but it could perhaps also refer to special devices that do not occur with adnominal nouns or verbs.

Perhaps the **gender-agreement suffixes** in Indo-European languages can be regarded as showing such “specialization”? For example, Modern Greek has

mikr-ó spiti ‘small house’
mikr-í póli ‘small town’

because *spíti* is neuter and *póli* is feminine.

But what about, say, German, which has very similar gender suffixes (*klein-es Haus* ‘small house’, *klein-e Stadt* ‘small town’),

but in addition has **agreeing participles** (e.g. *spiel-end-es Kind* ‘playing child’, *lach-end-e Frau* ‘laughing woman’)?

Should we say that in German, property-concept words **do not show a specialization for adnominal modification**, because verbs can be used in this way, too?

Thus:

it is not easy at all to define the term “**adjective**” formally in such a way that the criteria can be applied to all languages.

For this reason, I have proposed that **adjective** (as well as **verb** and **noun**) are defined notionally for comparative purposes (Haspelmath 2023).

“Adjectival predication” is best defined in functional terms, as **predication of a property concept**, not as a formally defined type of strategy (I use the term “**attributinal clause**”, adopted from Dixon 2010).

Admittedly, this goes against the most common view in the literature, which is reflected in statements such as the following:

“When a language has no adjectives or only a small closed class of adjectives, property concepts are **encoded either as verbs**, as in Chinese and Korean, **or nouns**, as in Quechua and Hausa (see Stassen 1997 for an overview).” (Nikolaeva & Spencer 2020: 7)

It is an extremely widespread view among linguists that some languages have a special adjective class, while others do not.

However, if “adjective” is **defined notionally**, then all languages have such a class, because all languages can express notions such as ‘big’, ‘small’, ‘old’, ‘good’ by simple roots.

Creissels et al. (2026: 10) also take this widespread view:

“Recognizing adjectival predication as a particular type of non-verbal predication makes only sense for languages with a class of adjectives whose predicative behavior is distinct from that of verbs.”

In this sentence, they say that the “predicative behaviour” should be “distinct”, in contrast to the earlier characterization of “adjectives” with respect to their adnominal function.

And what exactly is “distinct”?

Verb-like adjectives often lack the full range of tense-aspect forms that can occur with action verbs, but the occurrence of tense-aspect forms typically depends on actionality characteristic

(for example, progressives do not occur with stative verbs such as ‘know’ – does this make them “adjectives”?).

Thus, according to the view that I presented in Haspelmath (2025a), a clause such as (4a) (where *wouj* ‘red’ occurs with the past-tense prefix *té-*, like verbs), is attributinal, and counts as non-verbal, in contrast with the verbal clause in (4b).

(4) Haitian Creole:

- a. *Flè sa a té-wouj.*
[flower that DEF PST-red]
‘This flower was red.’
- b. *Chen sa a té-ale lakay.*
[dog that defpst-go home]
‘This dog had gone home.’

The term **nonverbal clause construction** is interpreted in functional terms: It refers to clause constructions in which there is **no action or process root** (= no *verb*, in the comparative-concept sense).

From a comparative sense, a copula or a ‘have’ verb in predpossession constructions is not a “verb”, though from language-particular perspectives, such elements are of course often labeled in this way.

Creissels et al. (2026), by contrast, want to **exclude clauses with property-concept roots that are too similar to action and process roots**, though they are not very specific (would they even exclude Japanese, where property-concept words such as *akai* ‘red’ are fairly verb-like?).

They also want to **exclude clauses with ‘have’ expressions that are called “verbs”** in the language, but again, they are not very specific (how verb-like does the ‘have’ word have to be?

Is English *have* sufficiently verb-like, even though it lacks an imperative and a passive construction?).

It may be that they are doing this because they are not fully aware of the distinction between **construction-functions** and **construction-strategies** (Haspelmath 2025b): Their notion of “nonverbal predication” is a kind of mixture of the two.

2. Word classes: definitions vs. universals

(Haspelmath 2023)

- the study of cross-linguistic word class variability and its limits needs to be based on clear comparative concepts because languages have different grammatical structures.
- the notions of (i) **function indicators** and (ii) **semantic root classes** give us a way of formulating universals and of distinguishing cross-linguistic macro-classes in a rigorous way.
- the root classes are action, property, and object; function indicators are markers that signal an unusual propositional act function (e.g. a copula signals predicative use of an object root, and a relativizer signals modifying use of an action root).
- we must distinguish clearly between comparison and language-particular analysis, because analysis must be based on language-particular constructions and is thus a rather different enterprise than the search for language universals.

3. Three word class universals

Word class universals are not particularly well-known, although many comparative linguists are aware of them and will not be surprised by the Universals 1-3 below.

Much of the comparative literature on word classes asks to what extent the **distinction** between nouns, verbs and adjectives is universal (e.g. Sasse 1993; Evans 2000; Baker 2003; Dixon 2010: Ch. 11). This question (which I refer to as the “**distinctness question**”) has no clear answer.

Universals 1-3 (due to Croft 1991: Ch. 3) are about the occurrence of **function indicators**, i.e.

copulas, attributivizers, and termifiers,

in three different **propositional act functions**:

predication, modification, and reference,

with three different **semantic root classes**:

action roots, property roots, and object roots.

Universal 1

If a language has a copula, i.e. a special form that indicates predicative function, it is used with object roots and/or property roots.

Universal 2

If a language has an attributivizer, i.e. a special form that indicates modifying function, it is used with action roots and/or object roots.

Universal 3

If a language has a termifier (= nominalizer or substantivizer), i.e. a special form that indicates referential function, it is used with property roots and/or action roots.

The reason the universals are formulated in terms of **roots**, not in terms of “words”, is that there is no general cross-linguistic understanding of what a “word” is (other than as defined by the conventional orthographic representation; Haspelmath 2011).

But in many or most languages, many roots can be words (especially nouns like *dog* and adjectives like *big*), and whenever a root cannot occur on its own without an affix (e.g. when tense affixes or person indexes are obligatory in verbs), these obligatory elements are not immediately relevant to Universals 1-3. Thus, formulating the universals in this way makes them generally applicable and we do not need to worry about “words”.

The construction types that the three universals make claims about are summarized in Table 1, where the boldfaced elements in the examples are the function indicators.

	reference:	modification:	predication:
	marked by termifiers	marked by attributivizers	marked by copulas
objects	–	genitive flag (4a: <i>the rent of the house</i>)	object-word copula (1a: <i>is a student</i>)
properties	substantivizer (5a: <i>the new one</i>)	–	property copula (2a: <i>is big</i>)
actions	nominalizer (6a: <i>the open-ing</i>)	relativizer (3a: <i>the work that they did</i>)	–

Table 1. Three propositional act functions and three semantic root classes

What the function indicators have in common is that they indicate an **unusual (unexpected, surprising) propositional act function**: A genitive flag indicates that an object root (unexpectedly, surprisingly) has modifying function, a property copula indicates that a property root has predicating function, a nominalizer indicates that an action root (unusually) has referential function, and so on.

4. The copula universal

Universal 1

If a language has a copula, i.e. a special form that indicates predicative function, it is used with object roots or property roots.

So we find many languages which require a copula with object-word predicates as in (1a), though not all languages have them (1b). Likewise, many languages require a copula with property predicates (2a), but some have no copula here (2b).

(1) a. Italian (copula *è*)

Michele è studente.
Michele is student
'Michele is a student.'

b. Russian (no copula)

Миша – студент.
Miša student.
Misha student
'Misha is a student.'

(2)a. English (copula *is*)

Our dog is big.

b. Mandarin Chinese (no copula)

我們的狗很大
Wǒmen de gǒu hěn dà.
we GEN dog very big
'Our dog is big.'

While a function indicator is commonly used for predicating object or property roots, it is not always necessary, so Universal 1 must be formulated as an implicational universal. But crucially, action roots generally do not require a form that indicates predicative function (in fact, they probably never do, as I do not now of a single example).

5. The attributivizer universal

Universal 2

If a language has an attributivizer, i.e. a special form that indicates modifying function, it is used with action roots or object roots.

The two main types of attributivizers are relativizers (relative markers, including participial affixes) and genitive flags (affixes or adpositions). So we find many languages which require a relativizer with action modifiers (= relative clauses) as in (3a), though not

all languages have them (3b). Likewise, many languages require a genitive flag with object-word modifiers (4a), but some have no marker here (4b).

(3)a. Lezgian (relativizer *-j*)

awu-nwa-j k'walax-ar

[do-PRF-REL] work-PL

‘the work that had been done’ (Haspelmath 1993: 344)

b. Japanese (no relativizer)

私がした仕事

watashi ga shi-ta shigoto

[I NOM do-PST] work

‘the work I did’

(4)a. Cape Verdean Creole (genitive flag *di* +)

kel renda di kaza

the rent of house

‘the rent of the house’ (Baptista 2013)

b. Seychelles Creole

lakaz sa zonm

house that man

‘that man’s house’ (Michaelis & Rosalie 2013)

So again, the function indicators do not occur universally, but property words are the least likely to require an attributivizer.

References

- Abdoulaye, Mahamane L. 2006. Existential and possessive predications in Hausa. *De Gruyter Mouton* 44(6). 1121–1164. (doi:10.1515/LING.2006.037)
- Beck, David. 2023. Adjectives. In van Lier, Eva (ed.), *The Oxford handbook of word classes*, 365–382. Oxford: Oxford University Press.
- Bentley, Delia & Ciconte, Francesco Maria & Cruschina, Silvio. 2015. *Existentials and locatives in Romance dialects of Italy*. Oxford: Oxford University Press.
- Bentley, Delia. 2017. Copular and existential constructions. In Dufter, Andreas & Stark, Elisabeth (eds.), *Manual of Romance morphosyntax and syntax*, 332–366. Berlin: De Gruyter. (doi 10.1515/9783110377088-009) (Accessed September 13, 2021.)
- Benveniste, Émile. 1960. «Etre» et «avoir» dans leurs fonctions linguistiques. *Bulletin de la Société de Linguistique de Paris* 55. (Also in: *Problèmes de linguistique générale* 1, 187–207. Paris: Gallimard, 1966)
- Bertinetto, Pier Marco & Ciucci, Luca & Creissels, Denis (eds.). 2026. *Non-verbal predication in the world’s languages: A typological survey (Volumes 1-2)*. Berlin: De Gruyter Mouton. (<https://www.degruyterbrill.com/document/doi/10.1515/9783110730982-001/html>)
- Bertinetto, Pier Marco & Ciucci, Luca & Farina, Margherita. 2019. Two types of morphologically expressed non-verbal predication. *Studies in Language*. John Benjamins 43(1). 120–194. (doi:10.1075/sl.17013.ber)
- Bickerton, Derek. 1981. *Roots of language*. Ann Arbor: Karoma Publishers [Republished in 2016 by Language Science Press: <https://langsci-press.org/catalog/book/91>].
- Borsley, Robert D & Tallerman, Maggie & Willis, David. 2007. *The syntax of Welsh*. Cambridge: Cambridge University Press.
- Chappell, Hilary & Lü, Shanshan. 2022. A semantic typology of location, existence, possession and copular verbs: Areal patterns of polysemy in Mainland East and Southeast Asia. *Linguistics* 60(1). 1–82. (doi:10.1515/ling-2021-0219)
- Clark, Eve V. 1978. Locational: Existential, locative, and possessive constructions. In Greenberg, Joseph H. (ed.), *Universals of human language, Vol. 4: Syntax*, 85–126. Stanford: Stanford University Press.

- Creissels, Denis & Bertinetto, Pier Marco & Ciucci, Luca. 2025. Non-verbal predication in the world's languages: An overview. In: *Oxford handbook of non-verbal predication* (to appear).
- Creissels, Denis & Bertinetto, Pier Marco & Ciucci, Luca. 2026. Non-verbal predication: An analytical framework. In Bertinetto, Pier Marco & Ciucci, Luca & Creissels, Denis (eds.), *Non-verbal predication in the world's languages: A typological survey (Volume 1)*, 1–56. Berlin: De Gruyter Mouton. (<https://www.degruyterbrill.com/document/doi/10.1515/9783110730982-001/html>)
- Creissels, Denis. 2013. Control and the evolution of possessive and existential constructions. In van Gelderen, Elly & Cennamo, Michela (eds.), *Argument Structure in Flux: the Naples-Capri papers*, 461–476. Amsterdam: Benjamins.
- Creissels, Denis. 2019. Inverse-locational predication in typological perspective. *Italian Journal of Linguistics* 31(2). 37–106.
- Creissels, Denis. 2020. La prédication existentielle et les constructions transpossessives. In Bottineau, Tatiana (ed.), *La prédication existentielle dans les langues naturelles : Valeurs et repérages, structures et modalités*, 17–41. Paris: Presses de l'Inalco.
- Croft, William. 2022. *Morphosyntax: Constructions of the world's languages*. Cambridge: Cambridge University Press. (<http://www.unm.edu/~wcroft/WACpubs.html>)
- Curnow, Timothy Jowan. 2000. Towards a cross-linguistic typology of copula constructions. In Henderson, John (ed.), *Proceedings of the 1999 Conference of the Australian Linguistic Society*, 1–9. (<http://www.als.asn.au>)
- Declerck, Renaat. 1983. 'It is Mr. Y' or 'he is Mr. Y'? *Lingua* 59(2). 209–246. (doi:10.1016/0024-3841(83)90064-5)
- Declerck, Renaat. 1988. *Studies on copular sentences, clefts and pseudo-clefts*. Leuven: Leuven University Press & Foris.
- den Dikken, Marcel & O'Neill, Teresa. 2017. Copular constructions in syntax. *Oxford Research Encyclopedia of Linguistics*. Oxford: Oxford University Press. (doi:10.1093/acrefore/9780199384655.013.137)
- Diessel, Holger. 1997. Predicative demonstratives. *Annual Meeting of the Berkeley Linguistics Society* 23(1). 72–82.
- Diessel, Holger. 1999. *Demonstratives: Form, function, and grammaticalization*. Amsterdam: Benjamins.
- Dik, Simon C. 1997. *The theory of Functional Grammar: 2 volumes*. Berlin: Mouton de Gruyter.
- Dixon, R. M. W. 2010. *Basic linguistic theory: Volume 1*. Oxford: Oxford University Press.
- Dixon, R. M. W. 2010. *Basic linguistic theory: Volume 2*. Oxford: Oxford University Press.
- Dryer, Matthew S. 2007. Clause types. In Shopen, Timothy (ed.), *Language typology and syntactic description, Vol. 1*, 224–275. Second Edition. Cambridge: Cambridge University Press.
- Eid, Mushira. 1991. Verbless sentences in Arabic and Hebrew. In Comrie, Bernard & Eid, Mushira (eds.), *Perspectives on Arabic Linguistics III: Papers from the Third Annual Symposium on Arabic Linguistics*, 31–61. Amsterdam: Benjamins.
- Evans, Nicholas & Levinson, Stephen C. 2009. The myth of language universals: Language diversity and its importance for cognitive science. *Behavioral and Brain Sciences* 32(5). 429–448. (doi:10.1017/S0140525X0999094X)
- Falk, Yehuda N. 2006. *Subjects and Universal Grammar: An explanatory theory*. Cambridge: Cambridge University Press.
- Farquharson, Joseph T. 2013. Jamaican structure dataset. In Michaelis, Susanne Maria & Maurer, Philippe & Haspelmath, Martin & Huber, Magnus (eds.), *Atlas of Pidgin and Creole Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. (<https://apics-online.info/contributions/8>)
- Foley, William A. & Van Valin, Robert D. 1977. On the viability of the notion of "subject" in universal grammar. *Berkeley Linguistics Society* 3. 293–320.
- Forker, Diana. 2013. *A grammar of Hinuq*. Berlin: De Gruyter Mouton.
- Francis, Elaine J. & Matthews, Stephen. 2005. A multi-dimensional approach to the category 'verb' in Cantonese. *Journal of Linguistics* 41(2). 269–305. (doi:10.1017/S0022226705003270)
- Freeze, Ray. 2001. Existential constructions. In Haspelmath, Martin & König, Ekkehard & Oesterreicher, Wulf & Raible, Wolfgang (eds.), *Language typology and language universals: An international handbook*, 941–953. Berlin: Mouton de Gruyter.
- Geist, Ljudmila. 2007. Predication and equation in copular sentences: Russian vs. English. In Comorovski, Ileana & von Heusinger, Klaus (eds.), *Existence: Semantics and syntax*, 79–105. Dordrecht: Springer.
- Gibson, Hannah & Guérois, Rozenn & Marten, Lutz. 2019. Variation in Bantu copula constructions. In Arche, Maria J. & Fábregas, Antonio & Marín, Rafael (eds.), *The grammar of copulas across languages*, 213–242. Oxford: Oxford University Press.
- Grimm, Nadine. 2021. *A grammar of Gyeli* (Comprehensive Grammar Library 2). Berlin: Language Science Press.
- Hamari, Arja. 2022. The negative existential cycle in Moksha Mordvin: From a negative existential into a negative auxiliary. In Veselinova, Ljuba & Hamari, Arja (eds.), *The Negative Existential Cycle*, 325–356. Berlin: Language Science Press. (DOI: 10.5281/zenodo.7353613)
- Haspelmath, Martin & Buchholz, Oda. 1998. Equative and similative constructions in the language of Europe. In Auwera, van der (ed.), *Adverbial constructions in the languages of Europe* (Empirical Approaches to Language Typology/EUROTYP, 20-3), 277–334. Berlin.
- Haspelmath, Martin. 1993. *A grammar of Lezgian* (Mouton Grammar Library 9). Berlin: Mouton de Gruyter.
- Haspelmath, Martin. 2011. On S, A, P, T, and R as comparative concepts for alignment typology. *Linguistic Typology* 15(3). 535–567.

- 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)
- Haspelmath, Martin. 2018. How comparative concepts and descriptive linguistic categories are different. In Van Olmen, Daniël & Mortelmans, Tanja & Brisard, Frank (eds.), *Aspects of linguistic variation: Studies in honor of Johan van der Auwera*, 83–113. Berlin: De Gruyter Mouton. (<https://zenodo.org/record/3519206>)
- Haspelmath, Martin. 2020a. The structural uniqueness of languages and the value of comparison for description. *Asian Languages and Linguistics* 1(2). 346–366.
- Haspelmath, Martin. 2020b. The morph as a minimal linguistic form. *Morphology* 30(2). 117–134. (doi:10.1007/s11525-020-09355-5)
- Haspelmath, Martin. 2023. Word class universals and language-particular analysis. In van Lier, Eva (ed.), *The Oxford handbook of word classes*, 15–40. Oxford: Oxford University Press.
- Haspelmath, Martin. 2024. Construction-functions and construction-strategies. To appear.
- Haspelmath, Martin. 2025a. Construction-functions and construction-strategies. In Däbritz, Chris Lasse & Budzisch, Josefine & Basile, Rodolfo (eds.), *Locative and existential predication: On forms, functions and neighboring domains*, 9–44. Berlin: Language Science Press. (doi: 10.5281/zenodo.16838056)
- Haspelmath, Martin. 2025b. Nonverbal clause constructions. *Language and Linguistics Compass* 19(2). e70007. (doi:10.1111/Inc3.70007)
- Hedberg, Nancy & Potter, David. 2010. Equative and predication copulas in Thai. *Berkeley Linguistics Society* 36(1). 144–157. (doi:10.3765/bls.v36i1.3908)
- Heine, Bernd. 1997. *Possession: Cognitive sources, forces, and grammaticalization*. Cambridge: Cambridge University Press.
- Hengeveld, Kees. 1992. *Non-verbal predication: Theory, typology, diachrony*. Berlin: Mouton de Gruyter.
- Heycock, Caroline & Kroch, Anthony. 1999. Pseudocleft connectedness: Implications for the LF Interface Level. *Linguistic Inquiry* 30(3). 365–397. (doi:10.1162/002438999554110)
- Heycock, Caroline. 2012. Specification, equation, and agreement in copular sentences. *Canadian Journal of Linguistics/Revue canadienne de linguistique*. Cambridge University Press 57(2). 209–240. (doi:10.1017/S0008413100004758)
- Higgins, F. R. 1973. *The pseudo-cleft construction in English*. Cambridge, MA: MIT dissertation. (<https://dspace.mit.edu/handle/1721.1/12988>)
- Higgins, F. R. 1979. *The pseudo-cleft construction in English*. New York: Garland.
- Huddleston, Rodney & Pullum, Geoffrey K. (eds.). 2002. *The Cambridge grammar of the English language*. Cambridge: Cambridge University Press.
- Izre'el, Shlomo. 2023. The syntax of presentative constructions: The view from colloquial Israeli Hebrew. In T.I. Davidjuk, I.I. Isaev, Ju.V. Mazurova, S. G. Tatevosov & O.I. Fedorova (eds.), *Jazyk kak on est': Sbornik statej k 60-letiju Andreja Aleksandroviča Kibrika*, 121–128. Moscow: Buki Vedi.
- Jespersen, Otto. 1924. *The philosophy of grammar*. London: Allen & Unwin.
- Killian, Don. 2021. Towards a typology of predicative demonstratives. *Linguistic Typology*. (doi:10.1515/lingty-2021-2078)
- Kim, Jong-Bok. 2016. Copular constructions and asymmetries in the specificational pseudocleft constructions in Korean. *Language and Linguistics*. SAGE Publications Ltd 17(1). 89–112. (doi:10.1177/1606822X15586672)
- Koch, Peter. 2012. Location, existence, and possession: A constructional-typological exploration. *Linguistics* 50(3). 533–603. (doi:10.1515/ling-2012-0018)
- Kornfilt, Jaklin. 1997. *Turkish*. London: Routledge.
- Kuno, Susumu & Wongkhamthong, Preya. 1981. Characterizational and Identificational Sentences in Thai. *Studies in Language* 5(1). 65–109. (doi:10.1075/sl.5.1.04kun)
- Lefebvre, Claire & Brousseau, Anne-Marie. 2002. *A grammar of Fongbe*. Berlin: Mouton de Gruyter.
- Locker, Ernst. 1954. Etre et avoir: Leurs expressions dans les langues. *Anthropos* 49(3/4). 481–510.
- Lyons, John. 1967. A note on possessive, existential and locative sentences. *Foundations of Language* 3(4). 390–396.
- Lyons, John. 1968. *Introduction to theoretical linguistics*. Cambridge: Cambridge University Press.
- Lyons, John. 1977. *Semantics*. Cambridge: Cambridge University Press.
- Maslova, Elena. 2003. *A grammar of Kolya Yukaghir* (Mouton Grammar Library 27). Berlin: Mouton de Gruyter.
- Matthews, Stephen & Yip, Virginia. 2011. *Cantonese: A comprehensive grammar*. London: Routledge.
- McGregor, William B. 2001. Non-verbal predicative possession in Nyulnyulan languages. In Simpson, Jane & Nash, David & Laughren, Mary & Austin, Peter & Alpher, Barry (eds.), *Forty years on: Ken Hale and Australian languages*, 337–352. Canberra: Australian National University. (Accessed September 13, 2021.)
- McNally, Louise. 2016. Existential sentences crosslinguistically: Variations in form and meaning. *Annual Review of Linguistics* 2(1). 211–231. (doi:10.1146/annurev-linguistics-011415-040837)
- Meillet, Antoine. 1906. La phrase nominale en indo-européen. *Mémoires de la Société de Linguistique de Paris* 14. 1–26.
- Michaelis, Susanne Maria & Rosalie, Marcel. 2013. Seychelles Creole structure dataset. In Michaelis, Susanne Maria & Maurer, Philippe & Haspelmath, Martin & Huber, Magnus (eds.), *Atlas of Pidgin and Creole Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. (<https://apics-online.info/contributions/56>)

- Mikkelsen, Line. 2005. *Copular clauses: Specification, predication and equation*. Amsterdam: Benjamins.
- Mikkelsen, Line. 2007. On so-called truncated clefts. In Geist, Ljudmila & Rothstein, Björn (eds.), *Kopulaverben und Kopulasätze: Intersprachliche und intrasprachliche Aspekte*, 47–68. Tübingen: Niemeyer.
- Mikkelsen, Line. 2011. Copular clauses. In Maienborn, Claudia & von Stechow, Klaus & Portner, Paul (eds.), *Semantics: An international handbook of natural language meaning, volume 2*, 1805–1829. Berlin: De Gruyter Mouton.
- Milsark, Gary L. 1977. Toward an explanation of certain peculiarities of the existential construction in English. *Linguistic Analysis* 3(2). 1–29.
- Mizuno, Shogo. 2024. Noun juxtaposition for predication, possession, and conjunction: Beyond ambiguity avoidance. *Linguistic Typology at the Crossroads* 4(2). 220–272. (doi:[10.6092/issn.2785-0943/18271](https://doi.org/10.6092/issn.2785-0943/18271))
- Moro, Andrea. 1997. *The raising of predicates: Predicative noun phrases and the theory of clause structure*. Cambridge: Cambridge University Press.
- Moro, Andrea. 2006. Existential sentences and expletive *there*. In Everaert, Martin & van Riemsdijk, Henk (eds.), *The Blackwell companion to syntax, volume II*, 210–236. Oxford: Blackwell.
- Myler, Neil. 2016. *Building and interpreting possession sentences*. Cambridge MA: MIT Press.
- Nikolaeva, Irina & Spencer, Andrew. 2020. *Mixed categories: The morphosyntax of noun modification*. Cambridge: Cambridge University Press.
- Nordlinger, Rachel & Sadler, Louisa. 2007. Verbless clauses: Revealing the structure within. In Zaenen, Annie E & Simpson, Jane & King, Tracy Holloway & Grimshaw, Jane & Maling, Joan & Manning, Chris (eds.), *Architectures, rules, and preferences: Variations on themes by Joan W. Bresnan*, 139–160. Stanford: CSLI Publications.
- Nordlinger, Rachel. 1998. *A grammar of Wambaya, Northern Australia*. Canberra: Pacific Linguistics. (<http://dx.doi.org/10.15144/PL-C140>)
- Overall, Simon E. & Vallejos, Rosa & Gildea, Spike (eds.). 2018. *Nonverbal predication in Amazonian languages*. Amsterdam: Benjamins.
- Overall, Simon E. 2018. Nonverbal predicates and copula constructions in Aguaruna (Chicham). In Overall, Simon E. & Vallejos, Rosa & Gildea, Spike (eds.), *Nonverbal predication in Amazonian languages*, 135–161. Amsterdam: Benjamins.
- Padučeva, Elena V. 1985. *Vyskazyvanie i ego sootnesennost' s dejstvitel'nost'ju: Referencial'nye aspekty semantiki mestoimenij*. Moskva: Nauka.
- Paducheva, Elena V. 2008. Locative and existential meaning of Russian быть. *Russian Linguistics* 32(3). 147–158.
- Panagiotidis, Phoevos. 2015. *Categorial features: A generative theory of word class categories*. Cambridge: Cambridge University Press.
- Partee, Barbara H. & Borschev, Vladimir. 2007. Existential sentences, BE, and the genitive of negation in Russian. In Comorovski, Ileana & von Stechow, Klaus (eds.), *Existence: Semantics and syntax*, 147–190. Dordrecht: Springer.
- Partee, Barbara H. 2010. Specificational Copular Sentences in Russian and English. *Oslo Studies in Language* 2(1). (doi:10.5617/osla.127) (<https://journals.uio.no/osla/article/view/127>) (Accessed July 19, 2021.)
- Payne, Thomas. 1997. *Describing morphosyntax: A guide for field linguists*. Cambridge: Cambridge University Press.
- Petrollino, Sara. 2019. Existential predication in Hamar. *Nordic Journal of African Studies* 28(4). 27–27.
- Pustet, Regina. 2003. *Copulas: Universals in the categorization of the lexicon*. Oxford: Oxford University Press.
- Rehg, Kenneth L. & Sohl, Damian G. 1981. *Ponapean reference grammar*. Honolulu: University of Hawaii Press.
- Rennison, John R. 1997. *Koromfe* (Descriptive Grammar Series). London: Routledge.
- Rothstein, Susan. 1995. Small clauses and copular constructions. In Cardinaletti, Anna (ed.), *Small clauses*, 27–48. New York: Academic Press. (doi:10.1163/9780585492209_003)
- Roy, Isabelle. 2013. *Nonverbal predication: Copular sentences at the syntax-semantics interface*. Oxford: Oxford University Press.
- Sabbagh, Joseph. 2009. Existential sentences in Tagalog. *Natural Language & Linguistic Theory* 27(4). 675–719. (doi:10.1007/s11049-009-9083-3)
- Sebeok, Thomas A. 1943. The equational sentence in Hungarian. *Language* 19(4). 320–327. (doi:10.2307/409935)
- Stassen, Leon. 1997. *Intransitive predication*. Oxford: Oxford University Press.
- Stassen, Leon. 1997. *Intransitive predication*. Oxford: Oxford University Press.
- Stassen, Leon. 2005. Predicative possession. In Haspelmath, Martin & Dryer, Matthew S. & Gil, David & Comrie, Bernard (eds.), *The world atlas of language structure*, 474–477. Oxford: Oxford University Press. (<http://wals.info/chapter/117>)
- Stassen, Leon. 2009. *Predicative possession*. Oxford: Oxford University Press.
- Stolz, Thomas & Levkovych, Nataliya. 2019. On belonging: Preliminary thoughts on the typology of belonging-constructions. In Johanson, Lars & Mazzitelli, Lidia Federica & Nevskaya, Irina (eds.), *Possession in languages of Europe and North and Central Asia*, 313–363. Amsterdam: Benjamins. (<https://benjamins.com/catalog/slcs.206>)
- Stolz, Thomas. 2001. To be with X is to have X: Comitatives, instrumentals, locative, and predicative possession. *Linguistics* 39(2). 321–350. (doi:10.1515/ling.2001.013)

- Tham, Shiao Wei. 2013. Possession as non-verbal predication. *Annual Meeting of the Berkeley Linguistics Society*, vol. 39, 302–316.
- Turunen, Rigina. 2010. *Nonverbal predication in Erzya: Studies on morphosyntactic variation and part of speech distinctions*. Helsinki: University of Helsinki. (Ph.D. dissertation.) (<https://helda.helsinki.fi/handle/10138/19652>) (Accessed September 13, 2021.)
- Vallejos, Rosa. 2016. *A grammar of Kukama-Kukamiria: A language from the Amazon*. Leiden: Brill.
- van Gijn, Erik. 2006. *A grammar of Yurakaré*. Nijmegen: Radboud University. (Ph.D. dissertation.) (http://webdoc.ubn.ru.nl/mono/g/gijn_e_van/gramofyu.pdf)
- Van Valin, Robert D., Jr. 2005. *Exploring the syntax-semantics interface*. Cambridge: Cambridge University Press.
- Verhaar, John W. M. (ed.). 1967-1972. *The verb “be” and its synonyms: Philosophical and grammatical studies (5 volumes, 1967-1972)*. Dordrecht: Reidel.
- Veselinova, Ljuba & Hamari, Arja (eds.). 2022. *The Negative Existential Cycle* (Research on Comparative Grammar). *Language Science Press* (Research on Comparative Grammar). Berlin: Language Science Press. (Accessed September 13, 2021.)
- Veselinova, Ljuba. 2013. Negative existentials: A cross-linguistic study. *Rivista di Linguistica* 25(1). 107–145.
- Ye, Jingting. 2020. Independent and dependent possessive person forms: Three universals. *Studies in Language* 44(2). 363–406. (doi:10.1075/sl.19020.ye)
- Ye, Jingting. 2021. *Property words and adjective subclasses in the world’s languages*. Leipzig: Leipzig University. (Ph D dissertation.)

Voice constructions:

4A. Causatives and anticausatives

MARTIN HASPELMATH
Max Planck Institute for Evolutionary Anthropology

1. Valency constructions and voice constructions

- (1) Verbs with different valencies in English
- Kim relaxed.*
 - Kim saw the dog.*
 - Kim looked at the dog.*
 - Kim relied on the dog.*
 - Kim showed the dog to Lee.*
- (2)
- | | |
|--------------|----------------------------|
| <i>relax</i> | ⟨V, 1[NOM]⟩ |
| <i>see</i> | ⟨V, 1[NOM], 2[ACC]⟩ |
| <i>look</i> | ⟨V, 1[NOM], 2[at]⟩ |
| <i>rely</i> | ⟨V, 1[NOM], 2[on]⟩ |
| <i>show</i> | ⟨V, 1[NOM], 2[ACC], 3[to]⟩ |
- (3) The valency of a verb is the set of argument positions that the verb takes together with their grammatical properties.

argument linking:

- (4) Russian
- Маша любит Мишу.
Maš-a ljubit Miš-u.
 Masha-NOM loves Misha-ACC
 ‘Masha loves Misha.’
 - Маша нравится Мише.
Maš-a nravitsja Miš-e.
 Masha-NOM pleases Misha-DAT
 ‘Misha likes Masha.’ (Lit. ‘Masha is pleasing to Misha.’)

The same argument type may correspond to very different roles: in (5), the dative argument *Miše* is not an experiencer, but a kind of beneficiary.

- (5) Саша помогает Мише.
Saš-a pomagaet Miš-e.
 Sasha-NOM helps Misha-DAT
 ‘Sasha helps Misha.’

Complete information about a verb must include not only its list of arguments with coding properties, but also its meaning and the linking of its participants to the arguments:

- (6) *ljubit'* $\langle V, 1[NOM], 2[ACC] \rangle$ '1 loves 2' (cf. 6a)
nravit'sja $\langle V, 1[NOM], 3[DAT] \rangle$ '3 likes 1' (cf. 6b)
pomogat' $\langle V, 1[NOM], 3[DAT] \rangle$ '1 helps 3' (cf. 7)

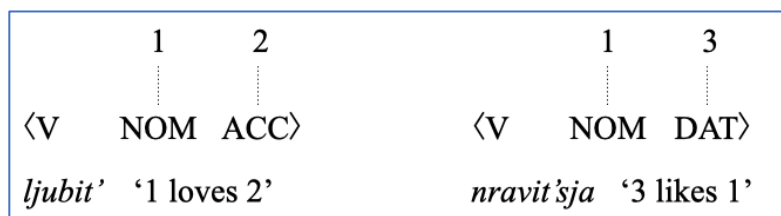


Figure 1: Linking diagrams for Russian *ljubit'* and *nravit'sja*

general valency construction (with variables X, Y, ...):

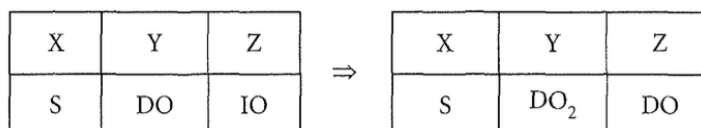
- (7) a. $\langle V, X[NOM] \rangle$ e.g. *relax, joke, work, die* (intransitive)
 b. $\langle V, X[NOM], Y[ACC] \rangle$ e.g. *see, kill, destroy, hate* (transitive)

flexivalency: (cf. Creissels 2024)

- (8) *subir* $\langle V, 1[NOM], 2[ACC] \rangle$ '1 makes 2 go up'
 \approx $\langle V, 2[NOM] \rangle$ '2 goes up'
- (9) *show* $\langle V, 1[NOM], 2[ACC], 3[to] \rangle$ '1 shows 2 to 3'
 \approx $\langle V, 1[NOM], 3[ACC], 2[ACC] \rangle$
- (10) English Labile alternation
 $\langle V, X[NOM], Y[ACC] \rangle$ 'X causes Y to undergo V'
 \approx $\langle V, Y[NOM] \rangle$ 'Y undergoes V'
- (11) English Dative alternation
 $\langle V, X[NOM], Y[ACC], Z[to] \rangle$ 'X causes Z to have/experience Y'
 \approx $\langle V, X[NOM], Z[ACC], Y[ACC] \rangle$

very similar to Kulikov (2011: 383):

(44) Dative shift (and dative passive)



This notation can also be used for “argument structure alternations” as discussed in the literature in English Construction Grammar (Goldberg 1995):

- (12) a. *She sneezed.*
 b. *She sneezed the napkin off the table.* (“argument structure construction”)

- (13) English Caused-Motion alternation
 $\langle V_\mu, X[NOM] \rangle$ ‘X acts in a μ manner’
 $\approx \langle V_\mu, X[NOM], Y[ACC], Z[DIRECTION] \rangle$ ‘X causes Y to move to Z by acting in a μ manner’

(Note: “Argument structure” and “valency” is basically the same.)

A **voice alternation** is any valency alternation with verbal coding (Zúñiga & Kittilä 2019; Creissels 2024), e.g. **passive alternation**

- (14) Swahili
 a. *Musa a-li-safisha nyumba.*
 Musa 3SG.HUM-PST-clean house.G7
 ‘Musa cleaned the house.’
 b. *Nyumba i-li-safish-wa na Musa.*
 house.G7 G7-PST-clean-PASS by Musa
 ‘The house was cleaned by Musa.’

- (15) Swahili Passive alternation
 $\langle [SBJ_x].[OBJ_y].V, X, Y \rangle$ ‘X acts on Y’
 $\approx \langle [SBJ_y].V\text{-}wa, Y, na+X \rangle$

anticausative alternations (Haspelmath 1987; based on Nedjalkov & Sil’nickij 1969)

- (5) Luragooli (Bantu; Gluckman & Bowler 2016: 272)

- a. *Sira a-hani muriango.*
 Sira 3SG.G1-close door(G3).
 ‘Sira closed the door.’

- b. *Muriango gu-han-iki.*
 door(G3) 3SG.G3-close-ANTC
 ‘The door closed.’

- (4) a. passive $\langle V, A_x, P_y \rangle$
 $\approx \langle V\text{-PASS}, \{obl\}_x, S_y \rangle$ ‘(X) acts on Y’
 d. anticausative $\langle V, A_x, P_y \rangle$ ‘X causes Y to change’
 $\approx \langle V\text{-ANTC}, S_y \rangle$ ‘Y changes’

Passive, antipassive, causative, and anticausative alternations:

- (16) a. passive $\langle V, A_X, P_Y \rangle$
 $\approx \langle V\text{-PASS}, \{\text{obl}\}_X, S_Y \rangle$ ‘(X) acts on Y’
- b. antipassive $\langle V, A_X, P_Y \rangle$
 $\approx \langle V\text{-ANTP}, S_X, \{\text{obl}\}_Y \rangle$ ‘X acts (on Y)’
- c. causative $\langle V, S_X \rangle$ ‘X acts’
 $\approx \langle V\text{-CAUS}, A_Z, P_X \rangle$ ‘Z makes X act’
- d. anticausative $\langle V, A_X, P_Y \rangle$ ‘X causes Y to change’
 $\approx \langle V\text{-ANTC}, S_Y \rangle$ ‘Y changes’

<p>passive voice alternation:</p> $\begin{array}{ccc} \langle V & A & P \rangle \\ & & \\ \langle V\text{-PASS} & \{\text{obl}\} & S \rangle \end{array}$	<p>antipassive voice alternation:</p> $\begin{array}{ccc} \langle V & A & P \rangle \\ & & \\ \langle V\text{-ANTP} & S & \{\text{obl}\} \rangle \end{array}$
<p>causative voice alternation:</p> $\begin{array}{ccc} \langle V & & S \rangle \\ & & \\ \langle V\text{-CAUS} & A & P \rangle \end{array}$	<p>anticausative voice alternation:</p> $\begin{array}{ccc} \langle V & A & P \rangle \\ & & \\ \langle V\text{-ANTC} & & S \rangle \end{array}$

Note: the term *voice* was traditionally restricted to passives and reflexives (plus perhaps antipassives), but more recently, it has been extended to causatives and applicatives (and other verb-coded valency alternations).

(Two important books: Zúñiga & Kittilä 2019; Creissels 2024.)

2. Two causative universals

2.1. Introduction to the two causative universals

(Haspelmath 2016)

Universal 1

If a language has synthetic causative marking for transitives, it will also have synthetic causative marking for intransitives (but not necessarily vice versa).

Universal 2

No language has labile causal-noncausal verb pairs with an unergative noncausal verb.

(2) a. Lezgian (Haspelmath 1993b: §10.1)

<i>c'ag-un</i>	freeze	<i>c'ag-ur-un</i>	make freeze	(intr.)
<i>gaton</i>	hit	<i>*gat-ur-un</i>	make hit	(tr.)

b. Makassarese (Jukes 2020: 288; 302)

<i>lari</i>	run	<i>pa-lari</i>	put to flight	(intr.)
<i>tinro</i>	sleep	<i>pa-tinro</i>	put to sleep	
<i>battu</i>	come	<i>pa-battu</i>	make come, bring	
<i>ta'gala'</i>	grasp	<i>pi-ta'gal-i</i>	make/let s.o grasp s.t	(tr.)
<i>erang</i>	carry	<i>pi-erangng-i</i>	make/let s.o carry/bear s.t	
<i>balli</i>	buy	<i>pi-ball-i</i>	make/let s.o buy s.t	

(3) (Hale 2000: 159)

a. labile verbs in English

<i>The water boiled.</i>	<i>We boiled the water.</i>
<i>The shirt dried.</i>	<i>The sun dried the shirt.</i>
<i>The tank filled.</i>	<i>The attendant filled the tank.</i>

b. an unattested alternation type

<i>The baby cried.</i>	<i>*The noise cried the baby.</i>
<i>The man coughed.</i>	<i>*The dust coughed the man.</i>
<i>The children played outdoors.</i>	<i>*The parents played the children outdoors.</i>

(OK: *The noise made the baby cry; The dust made the man cough;*
The parents made the children play outdoors.)

2.2. Efficiency explanations in a nutshell

- transitive verb roots (e.g. ‘hit’) occur with causal meaning added less commonly than intransitive verb roots, so this meaning tends to be **expressed analytically**
- unergative verb roots (e.g. ‘play’) occur with causal meaning added less commonly than unaccusative verb roots, so they are almost **never expressed without marking**
- passives involve extra coding, which is motivated if it signals **deviation from the usual associations of roles and referential prominence**, i.e. if the agent is less topical and/or the patient is more topical (Haspelmath 2021: 155)

3. Comparative concepts:

Causatives and anticausatives, causal and noncausal meanings

We need to distinguish between *causative/anticausative* as terms for strategies, and *causal/noncausal* as terms for types of verbal meanings:

Causal-noncausal verb pairs:

(formerly: “causative-inchoative”)

	<u>causal</u>	<u>noncausal</u>	(formal strategies:)
English	<i>break</i> (tr.)	<i>break</i> (intr.)	labile pair
Indonesian	<i>me-matah-kan</i>	<i>patah</i>	CAUSATIVE/basic pairs
Russian	<i>slomat'</i>	<i>slomat'-sja</i>	basic/ANTICAUSATIVE pair
Lithuanian	<i>laužti</i>	<i>lūžti</i>	equipollent pair

Causal-noncausal verb meanings may also be expressed analytically/syntactically:

<u>causal</u>	<u>noncausal</u>
<i>make so. laugh</i>	<i>laugh</i>
<i>make so. buy sth.</i>	<i>buy sth.</i>
<i>wash</i>	<i>undergo washing, be wash-ed</i>
<i>cut</i>	<i>undergo cutting, be cut</i>

These are COMPARATIVE CONCEPTS (Haspelmath 2010; 2018), and nothing is implied about the best analysis of such forms in particular languages.

Since Haspelmath (1993), there has been a lot of cross-linguistic research on such “transitivity pairs”, e.g. NINJAL’s *World Atlas of Transitivity Pairs*
(<https://watp.ninjal.ac.jp/en>)



We might simply refer to the two members of these pairs as “intransitive” and “transitive”, but there are other types of transitivity pairs (passive-nonpassive, antipassive-nonantipassive, etc.).

Thus, it is better to use a new term pair: *non-causal* vs. *causal* (Haspelmath 2016), reserving *anticausative* and *causative* for specific strategies

Shibatani (2016) still uses the old terminology:

Noncausative V		Causative V	
ak-u	open (vi)	ak-e-ru	open (vt)
akaram-u	redde (vi)	akaram-e-ru	let redde

Noncausative V		Causative V	
egur-e-ru	get scooped out	egur-u	scoop out
hag-e-ru	peel off (vi)	hag-u	tear off (vt)

The new terminology (*noncausal* vs. *causal*) has been widely adopted since 2016.

However, especially in works on European languages, causal-noncausal alternations are often called “causative-anticausative alternations” (e.g. Alexiadou et al. 2006).

(See Haspelmath et al. (2014: 590) for some discussion of the terminological issues; the term *anticausative* was originally introduced by Nedjalkov & Sil’nickij 1969.)

The marking of the causal-noncausal contrast can occur on the causal member (e.g. *make* so. *laugh*), or on the noncausal member (e.g. *be wash-ed*), or there is no marking at all (*break/break*)

– this means that this is a situation of **differential coding**:

the **same** meaning is coded **differently** under **different** conditions

(somewhat similar to differential object marking)

4. Anticausatives are paradoxical

Mel’čuk (1967) first pointed out that anticausatives violate a general principle of meaning-form correspondence.

causative	‘x MAKE (y BECOME open)’	<i>ak-e(ru)</i>
	‘y BECOME open’	<i>ak(u)</i>
anticausative	‘y BECOME open’	<i>hirak-e(ru)</i>
	‘x MAKE (y BECOME open)’	<i>hirak(u)</i>

In anticausatives, there is an extra formal element for a meaning that lacks an extra meaning element – this is the *opposite* of what one might expect.

This seems to violate a general functional principle of **iconicity**:

“more meaning corresponds to more form” (e.g. Haiman 1983, Givón 1991)

In Haspelmath (1993), I noted this problem, but did not have a good solution – I did not want to abandon the iconicity principle, **because I believed in it**.

It was only much later that I realized that this principle was **wrong** (see also this blogpost: <https://dlc.hypotheses.org/2873>).

In general, linguistic expressions are **longer when they convey rarely expressed meanings**, not when they express “complex meanings”.

This also applies to word length, e.g.

<i>mouse</i>	vs.	<i>crocodile</i>
<i>star</i>	vs.	<i>planet</i>
<i>phone</i>	vs.	<i>computer</i>

This is because languages have a tendency to exhibit *efficiency of coding* (or *communicative efficiency*, e.g. Levshina 2023).

5. Universals of causative and anticausative verb formation are probably explained by frequency of use

Some data from English:

(from the BNC; boldface percentages from Wright 2001:127-28):

	causative		noncausative	
agent-requiring	<i>cut</i>	100%	<i>undergo cutting</i>	0%
costly	<i>break (tr.)</i>	90%	<i>break (intr.)</i>	10%
	<i>open (tr.)</i>	80%	<i>open (intr.)</i>	20%
	<i>split (tr.)</i>	78%	<i>split (intr.)</i>	22%
(intermediate)	<i>burn (tr.)</i>	76%	<i>burn (intr.)</i>	24%
automatic	<i>melt (tr.)</i>	72%	<i>melt (intr.)</i>	28%
	<i>freeze (tr.)</i>	62%	<i>freeze (intr.)</i>	38%
	<i>dry (tr.)</i>	61%	<i>dry (intr.)</i>	39%
	<i>sink (tr.)</i>	58%	<i>sink (intr.)</i>	42%
(internally caused)	<i>sprout (tr.)</i>	33%	<i>sprout (intr.)</i>	67%
	<i>rot (tr.)</i>	14%	<i>rot (intr.)</i>	86%
	<i>rust (tr.)</i>	5%	<i>rust (intr.)</i>	95%
agentive intransitive	<i>make laugh</i>	20.8%	<i>laugh</i>	...
	<i>make cry</i>	6.0%	<i>cry</i>	
	<i>make jump</i>	5.6%	<i>jump</i>	
	<i>make vomit</i>	2.6%	<i>vomit</i>	
	<i>make dance</i>	1.8%	<i>dance</i>	
	<i>make weep</i>	1.1%	<i>weep</i>	
	<i>make sing</i>	0.4%	<i>sing</i>	
	<i>make wait</i>	0.4%	<i>wait</i>	
transitive	<i>make throw</i>	0.13%	<i>throw</i>	
	<i>make buy</i>	0.11%	<i>buy</i>	
	<i>make accept</i>	0.05%	<i>accept</i>	
	<i>make avoid</i>	0.05%	<i>avoid</i>	
	<i>make hit</i>	0.01%	<i>hit</i>	
	<i>make destroy</i>	0.00%	<i>destroy</i>	100%
	<i>make teach</i>	0.00%	<i>teach</i>	100%

Table 3.

Shibatani (2016): in some verb pairs, the causal is more frequent, while in others, the noncausal is more frequent

(46)	-Ø (Transitive)		-e (Intransitive)	
a.	<i>kir-u</i> 'cut'	46,000,000	<i>kir-e-ru</i>	25,200,000
	<i>yak-u</i> 'burn'	23,200,000	<i>yak-e-ru</i>	5,640,000
	<i>hag-u</i> 'peel off'	22,100,000	<i>hag-e-ru</i>	603,000
	<i>yabur-u</i> 'tear off'	7,070,000	<i>yabur-e-ru</i>	2,500,000
	<i>war-u</i> 'break'	7,940,000	<i>war-e-ru</i>	6,390,000
(47)	-Ø (Intransitive)		-e (Transitive)	
a.	<i>ak-u</i> 'open'	373,000,000	<i>ak-e-ru</i>	32,200,000
	<i>ukabu</i> 'float'	22,000,000	<i>ukab-e-ru</i>	3,300,000
	<i>sizum-u</i> 'sink'	10,100,000	<i>sizum-e-ru</i>	1,750,000
	<i>tizim-u</i> 'shrink'	2,870,000	<i>tizim-e-ru</i>	1,350,000
	<i>muk-u</i> 'face'	17,900,000	<i>muk-e-ru</i>	11,900,000

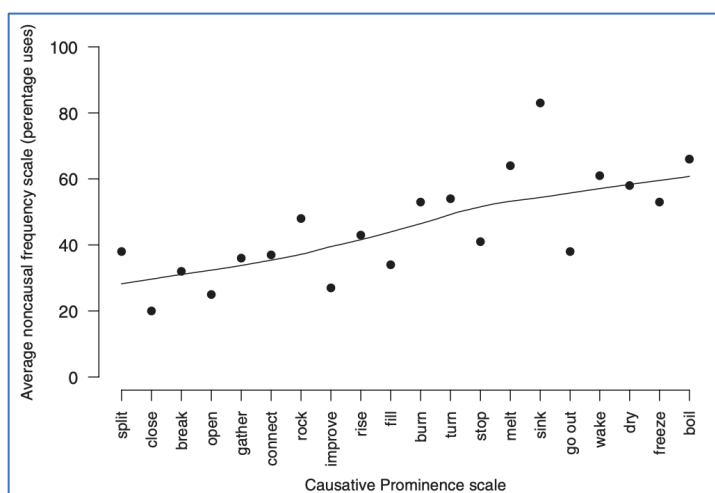
However, these figures cannot serve as an immediate explanation of the asymmetric coding in Japanese:

- the usage frequency cannot *directly cause* the coding
- there are exceptions, as noted by Shibatani, e.g.

b.	<i>nuk-u</i> 'pull out'	25,500,000	<i>nuk-e-ru</i>	30,400,000
	<i>or-u</i> 'break off'	4,300,000	<i>or-e-ru</i>	6,160,000
	<i>tur-u</i> 'fish'	7,880,000	<i>tur-e-ru</i>	9,470,000
b.	<i>tuk-u</i> 'attach'	43,900,000	<i>tuk-e-ru</i>	75,000,000
	<i>sodat-u</i> 'grow'	15,300,000	<i>sodat-e-ru</i>	32,400,000
	<i>kagam-u</i> 'bend down'	162,000	<i>kagam-e-ru</i>	309,000

Frequency of use is a *soft causal force* – overall, language systems tend to be efficient, i.e. more frequently occurring meanings are expressed by shorter forms. But this need not be manifested in every single instance. There are many local inefficiencies.

The trend is most clearly visible when we examine many languages from around the world (as I did in 1993), and when we look at corpus frequencies from six different languages:



6. A biocognitive explanation of the universal about labile unergatives: Hale (2000)

Universal 2

No language has labile causal-noncausal verb pairs with an unergative noncausal verb.

(e.g. **The noise cried the baby.*)

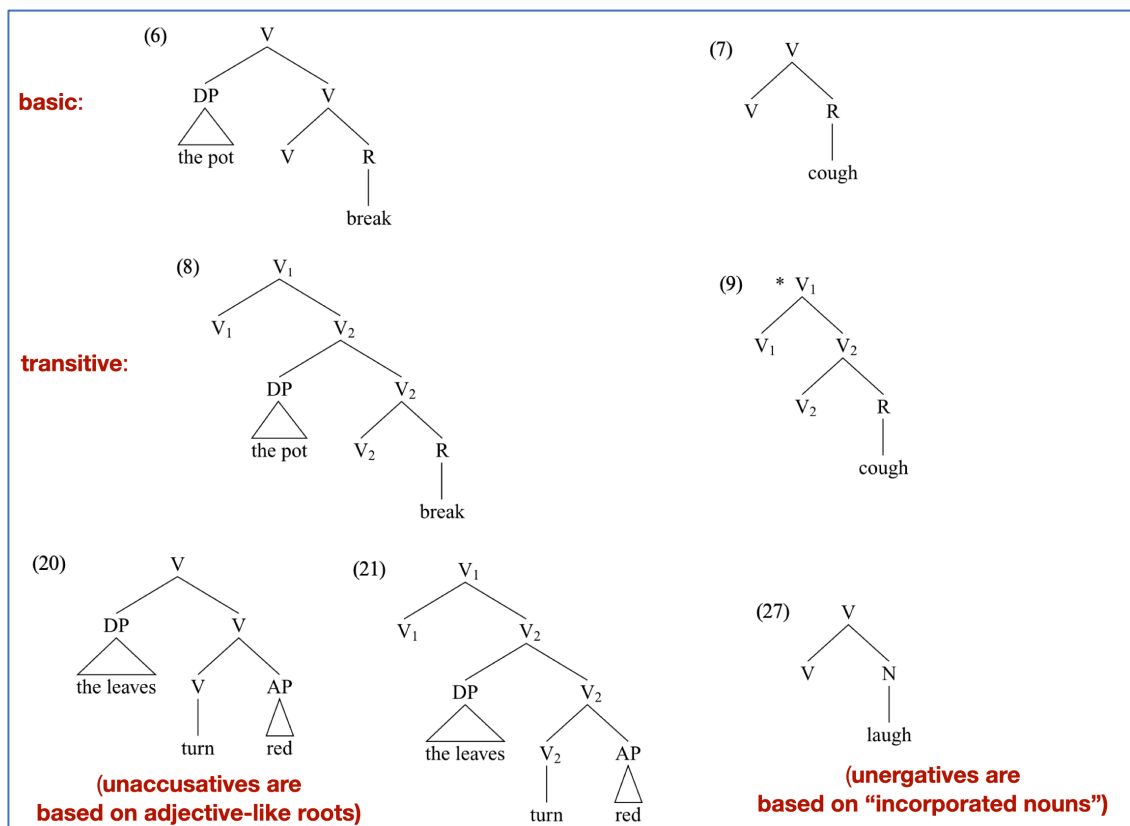
Hale (2000) and Hale & Keyer (2002: §4.4) propose that there are just four different ways in which heads can combine with complements and specifiers.

Crucially:

- adjectives have a **specifier**, but nouns **lack a specifier**
- unaccusatives have **adjective-like roots**,
unergatives have **noun-like roots**

We believe that the observations made here reflect general principles that govern the limits on lexical argument structures. The contrast noted here is the same, in essence, as a contrast that is virtually universal among the languages of the world. Languages differ in detail, but the basic elements and principles are the same.

(Hale & Keyer 2002: 135)



This explanation works only if it is a fixed condition of UG that adjectives and unaccusatives must have a specifier, while nouns and unergatives must lack a specifier.

This is possible but unlikely.

7. A differential-coding scale for causatives and anticausatives: The spontaneity scale

Differential coding can be stated not only in pairs of contexts, but also in terms of implicational scales (Haspelmath 2019):

place name	common noun	human noun
← less locative coding		more locative coding →

Basque:

Gasteiz-en
‘in Gasteiz’

zuhaitz-ean
‘on the tree’

mutila-gan
‘at the boy(‘s)’

Italian:

a Roma, a-lla chiesa
‘to Rome, to the church’

da-l poliziotto
‘to the policeman’

French:

Ø rue Molière
‘in Rue Molière’

dans le jardin, chez Marie
‘in the garden, at Marie’s’

In a similar way, differential coding for causal and noncausal verb meanings can be stated in a very general way by an implicational scale (Haspelmath 2016):

transitive	unergative	automatic	costly	agentful
		unaccusative		
← more causative coding				more anticausative coding →

English

make buy, make laugh

freeze, break

be washed

French

faire acheter, faire rire, faire fondre

se casser

être lavés

Makassarese

pi-ball-i ‘make buy’, *pa-lari* ‘m. run’, *pa-battu*
‘bring’

Indonesian

me-mecah-kan ‘break (tr.)’, *me-mandi-kan* ‘bathe (tr.)’

Russian

kipjat-it ‘make
boil’

rasplavit-sja ‘melt’, *otkryt-sja* ‘open’, *myt-sja* ‘be
washed’

Where does the distinction between the automatic ('freeze') and costly ('break') types come from?

It was inferred from cross-linguistic data (Haspelmath 1993a):

Table 4. Expression types by verb pairs

	total	A	C	E	L	S	A/C
18. 'boil'	21	0.5	11.5	3	6	0	0.04
25. 'freeze'	21	2	12	3	4	0	0.17
29. 'dry'	20	3	10	4	3	0	0.30
1. 'wake up'	21	3	9	6	2	1	0.33
20. 'go out/put out'	21	3	7.5	5.5	3	2	0.41
11. 'sink'	21	4	9.5	5.5	1.5	0.5	0.42
8. 'learn/teach'	21	3.5	7.5	6	2	3	0.47
13. 'melt'	21	5	10.5	3	2.5	0	0.48
31. 'stop'	21	5.5	9	3.5	3	0	0.61
23. 'turn'	21	8	7.5	4	1.5	0	1.07
26. 'dissolve'	21	10.5	7.5	2	1	0	1.40
3. 'burn'	21	7	5	2	5	2	1.40
14. 'destroy'	20	8.5	5.5	5	1	0	1.55
27. 'fill'	21	8	5	5	3	0	1.60
22. 'finish'	21	7.5	4.5	5	4	0	1.67
7. 'begin'	19	5	3	3	8	0	1.67
10. 'spread'	21	11	6	3	1	0	1.83
24. 'roll'	21	8.5	4.5	5	3	0	1.89
16. 'develop'	21	10	5	5	1	0	2.00
15. 'get lost/lose'	21	11.5	4.5	4.5	0	0.5	2.56
21. 'rise/raise'	21	12	4.5	3.5	0	1	2.67
28. 'improve'	21	8.5	3	8	1.5	0	2.67
19. 'rock'	21	12	40	3.5	1.5	0	3.00
17. 'connect'	21	15	2.5	1.5	1	1	6.00
12. 'change'	21	11	1.5	4.5	4	0	7.33
9. 'gather'	21	15	2	3	1	0	7.50
5. 'open'	21	13	1.5	4	2.5	0	8.67
2. 'break'	21	12.5	1	4	3.5	0	12.50
6. 'close'	21	15.5	1	2.5	2	0	15.50
30. 'split'	20	11.5	0.5	5	3	0	23.00
4. 'die/kill'	21	0	3	1	1	16	—
total	636	243	164.5	128.5	69	31	

Haspelmath (1993a) was inspired by Mel'čuk (1967), but at the time, I found my findings puzzling.

Levin & Rappaport Hovav (1995) took them up, but were puzzled, too.

(T. 1187)

In a follow-up to Nedjalkov's (1969) study discussed in section 3.1, Haspelmath (1993) discusses verbs that tend not to show consistent patterns of behavior across languages. For example, the morphologically simple form of the verb corresponding to English *melt* tends to be transitive in most languages, the intransitive form being the morphologically

derived form, but a few languages show the opposite pattern. It is likely that this cross-linguistic variation arises because the meaning of a verb such as *melt* is consistent with its describing either an internally or an externally caused eventuality. In fact, it should be possible to verify this

Explaining the scale generalizations in terms of frequency of use:

The proposal here is that the coding tendencies in the world's languages are explained by usage tendencies:

The base verb meanings to the left on the scale are least likely to occur with a causal meaning, while the likelihood of a causal meaning increases gradually as we go to the right of the scale.

References

- Akkuş, Faruk. 2021. *Argument introduction, Voice and causatives*. Philadelphia: University of Pennsylvania. (PhD dissertation.)
- Bahrt, Nicklas N. 2021. *Voice syncretism*. Berlin: Language Science Press. (<https://langsci-press.org/catalog/view/315/3359/2373-1>)
- Bhatt, Rajesh & Embick, David. 2017. Causative derivations in Hindi-Urdu. *Indian Linguistics* 78(1–2). 93–151.
- Cobbinah, Alexander & Lüpke, Friederike. 2012. Not cut to fit: Zero coded passives in African languages. In Brenzinger, Matthias & Fehn, Anne-Maria (eds.), *Proceedings of the 6th World Congress of African Linguistics*, 153–165. Cologne: Köppe.
- Comrie, Bernard. 1988. Passive and voice. In Shibatani, Masayoshi (ed.), *Passive and voice*, vol. 1, 9–23. Amsterdam: Benjamins.
- Cristofaro, Sonia. 2010. Semantic maps and mental representation. *Linguistic Discovery* 8(1). 35–52. (doi:10.1349/PS1.1537-0852.A.345)
- Croft, William & Shyldkrot, Hava Bat-Zeev & Kemmer, Suzanne. 1987. Diachronic semantic processes in the middle voice. In Giacalone Ramat, Anna & Carruba, Onofrio & Bernini, Giuliano (eds.), *Papers from the 7th International Conference on Historical Linguistics*. Amsterdam: Benjamins.
- Georgakopoulos, Thanasis & Polis, Stéphane. 2018. The semantic map model: State of the art and future avenues for linguistic research. *Language and Linguistics Compass* 12(2). e12270. (doi:10.1111/lnc3.12270)
- Hale, Kenneth L. & Keyser, Samuel Jay. 2002. *Prolegomenon to a theory of argument structure* (Linguistic Inquiry Monographs). Cambridge, Mass.: MIT Press. (<http://catdir.loc.gov/catdir/toc/fy037/2002024420.html>) (Accessed July 6, 2023.)
- Hale, Kenneth L. 2000. A Uto-Aztecan ('O'odham) reflection of a general limit on predicate argument structure. In Casad, Eugene H. & Willett, Thomas L. (eds.), *Uto-Aztecan: Structural, temporal and geographic perspectives: Papers in memory of Wick R. Miller by the friends of Uto-Aztecan*, 155–169. Hermosillo: Universidad de Sonora.
- Haspelmath, Martin. 1987. *Transitivity alternations of the anticausative type* (Arbeitspapiere Des Instituts Für Sprachwissenschaft N.F. Nr. 4). Cologne: Universität zu Köln. (<http://doi.org/10.5281/zenodo.225196>)
- Haspelmath, Martin. 1990. The grammaticization of passive morphology. *Studies in Language* 14(1). 25–72.
- Haspelmath, Martin. 1993. *A grammar of Lezgian* (Mouton Grammar Library 9). Berlin: Mouton de Gruyter.
- Haspelmath, Martin. 2003. The geometry of grammatical meaning: Semantic maps and crosslinguistic comparison. In Tomasello, Michael (ed.), *The New Psychology of Language*, vol. 2, 211–243. New York: Lawrence Erlbaum. (10.5281/zenodo.831410)
- Haspelmath, Martin. 2010. Comparative concepts and descriptive categories in crosslinguistic studies. *Language* 86(3). 663–687. (doi:10.1353/lan.2010.0021)
- Haspelmath, Martin. 2011. On S, A, P, T, and R as comparative concepts for alignment typology. *Linguistic Typology* 15(3). 535–567.
- 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)
- Haspelmath, Martin. 2018. How comparative concepts and descriptive linguistic categories are different. In Van Olmen, Daniël & Mortelmans, Tanja & Brisard, Frank (eds.), *Aspects of linguistic variation: Studies in honor of Johan van der Auwera*, 83–113. Berlin: De Gruyter Mouton. (<https://zenodo.org/record/3519206>)
- Haspelmath, Martin. 2021. Role-reference associations and the explanation of argument coding splits. *Linguistics* 59(1). 123–174. (doi:10.1515/ling-2020-0252)
- Haspelmath, Martin. 2022. Valency and voice constructions. Leipzig, MPI-EVA, ms. (<https://ling.auf.net/lingbuzz/005941>)

- Haspelmath, Martin. 2023. Coexpression and synexpression patterns across languages: Comparative concepts and possible explanations. *Frontiers in Psychology* (to appear). (<https://ling.auf.net/lingbuzz/007402>)
- Haspelmath, Martin. 2023a. Inflection and derivation are merely ways of speaking. (to appear).
- Jukes, Anthony. 2019. *A grammar of Makasar: A language of South Sulawesi, Indonesia* (Grammars and Sketches of the World's Languages). Boston: BRILL.
- Keenan, Edward L. & Dryer, Matthew S. 2007. Passives in the world's languages. *Language Typology and Syntactic Description. Vol. I: Clause Structure*, 325–361.
- Legate, Julie Anne & Akkuş, Faruk & Šereikaitė, Milena & Ringe, Don. 2020. On passives of passives. *Language* 96(4). 771–818. (doi:[10.1353/lan.2020.0062](https://doi.org/10.1353/lan.2020.0062))
- Legate, Julie Anne. 2021. Noncanonical passives: A typology of voices in an impoverished Universal Grammar. *Annual Review of Linguistics* 7(1). (doi:[10.1146/annurev-linguistics-031920-114459](https://doi.org/10.1146/annurev-linguistics-031920-114459))
- Levin, Beth & Rappaport Hovav, Malka. 1995. *Unaccusativity: At the syntax-lexical semantics interface*. Cambridge, MA: MIT Press.
- Mel'čuk, Igor. 1967. K ponjatiju slovoobrazovanija. *Izvestija AN SSSR: Serija literatury i jazyka* 26(4). 352–362.
- Neu, Eva. 2023. Against a low subject analysis of causatives of unergatives. *Proceedings of the Linguistic Society of America* 8(1). 5527. (doi:[10.3765/plsa.v8i1.5527](https://doi.org/10.3765/plsa.v8i1.5527))
- Perlmutter, David M. & Postal, Paul M. 1977. Toward a universal characterization of passivization. *Annual Meeting of the Berkeley Linguistics Society*, vol. 3, 394–417.
- Siewierska, Anna & Bakker, Dik. 2013. Passive agents: Prototypical vs. canonical passives. In Brown, Dunstan & Chumakina, Marina & Corbett, Greville G. (eds.), *Canonical morphology and syntax*, 151–189. Oxford: Oxford University Press.
- Siewierska, Anna. 1984. *The passive: A contrastive linguistic analysis*. London: Croom Helm.
- Siewierska, Anna. 2005. Passive constructions. In Haspelmath, Martin & Dryer, Matthew S. & Gil, David & Comrie, Bernard (eds.), *The world atlas of language structures*, 434–437. Oxford: Oxford University Press.
- Simpson, Adrian P. 1999. Fundamental problems in comparative phonetics and phonology: Does UPSID help to solve them. *Proceedings of the 14th international congress of phonetic sciences*, vol. 1, 349–352.
- Tollan, Rebecca. 2018. Unergatives are different: Two types of transitivity in Samoan. *Glossa: A Journal of General Linguistics* 3(1). (doi:[10.5334/gjgl.223](https://doi.org/10.5334/gjgl.223)) (<https://www.glossa-journal.org/article/id/5003/>) (Accessed July 6, 2023.)
- Tsunoda, Tasaku. 2011. *A grammar of Warrongo*. Berlin: De Gruyter Mouton. (<https://nbn-resolving.org/urn:nbn:de:101:1-201611304711>)
- Wichmann, Søren. 2009. Case relations in a head-marking language: Verb-marked cases in Tlapanec. In Malchukov, Andrej & Spencer, Andrew (eds.), *The Oxford handbook of case*, 797–807. Oxford: Oxford University Press.
- Wright, Sandra K. 2001. *Internally caused and externally caused change of state verbs*. Northwestern University. (Ph.D. dissertation.)

Nagoya 2024

1. Overview

- many languages have **causative** verbs (transitive),
such as Japanese *yom-ase(-ru)* ‘make someone read’ (§2)
- some languages also have **anticausative** verbs (intransitive),
such as Russian *otkryt’-sja* ‘(become) open’ (compare *otkryt’* ‘(make) open’) (§3)

- transitivity pairs may also be expressed by **labile** verbs, e.g. English *to open*
(*the door opened* – *we opened the door*) (§4)
- it often seems arbitrary which coding type a language chooses – but we can try to find some general patterns (Nedjalkov & Sil'nickij 1969; Shibatani (ed.) 1976; Haspelmath 1993; 2016)
- two competing hypotheses:
 - iconicity (meaning-form correspondence; Haspelmath 1993)
 - efficiency of coding (form-frequency correspondence; Haspelmath 2016)
- the cross-linguistic evidence favours the efficiency hypothesis, and more generally, a social-functional theory of language structures

2. Causative verb formation

Some examples from Japanese (from Shibatani 2016: 479):

Noncausative V		Causative V	
ak-u	open (vi)	ak-e-ru	open (vt)
akaram-u	reddden (vi)	akaram-e-ru	let reddden
dok-u	get out of the way	dok-e-ru	put out of the way
hisom-u	lurk	hisom-e-ru	hide (self)
hukum-u	include (in self)	hukum-e-ru	include (in another)
hus-u	lie down	hus-e-ru	lay down
itam-u	hurt	itam-e-ru	injure
kagam-u	stoop, get bent	kagam-e-ru	bend (one's body)
kanasim-u	grieve	kanasim-e-ru	make grieve
kanaw-u	come true	kanaw-e-ru	make come true

Some examples from other languages (cited from Haspelmath 2016: 41):

Tuvan (Kulikov 1998a: 260)

- (11) a. *ool doŋ-gan*
 boy freeze-PST
 'The boy froze.'
- b. *ašak ool-du doŋ-ur-gan*
 old.man boy-ACC freeze-CAUS-PST
 'The old man made the boy freeze.'
- (12) a. *ašak ool-du ette-en*
 old.man boy-ACC hit-PST
 'The old man hit the boy.'
- b. *Bajır ašak-ka ool-du ette-t-ken*
 Bajır old.man-DAT boy-ACC hit-CAUS-PST
 'Bajır made the old man hit the boy.'

Indonesian (Cole & Son 2004: examples 1, 2, 5)

- (13) a. *Cangkir-nya pecah.*
 cup-DEF break
 'The cup broke.'
- b. *Tono me-mecah-kan cangkir-nya.*
 Tono ACT-break-CAUS cup-3
 'Tono broke the cup.'

3. Anticausative verb formation

Some examples from Japanese (from Shibatani 2016: 484):

Noncausative V		Causative V	
egur-e-ru	get scooped out	egur-u	scoop out
hag-e-ru	peel off (vi)	hag-u	tear off (vt)
hirak-e-ru	become open & wide	hirak-u	open (vt)
hodok-e-ru	come untied	hodok-u	untie
hur-e-ru	swing (vi)	hur-u	shake (vt)
kak-e-ru	become lacking	kak-u	lack (vt)
kasur-e-ru	get scratchy/hoarse	kasur-u	graze, brush
kezur-e-ru	get shaved/planed	kezur-u	shave, plane
kir-e-ru	get cut off/severed	kir-u	cut, sever
kosur-e-ru	rub (vi)	kosur-u	rub, scrape

Some examples from other languages (cited from Haspelmath 2016: 41):

(4)	basic verb (causal meaning)	anticausative verb (with special coding) (noncausal meaning)
a. Russian	<i>otkryt</i> ‘open (tr.)’	<i>otkryt’-sja</i> ‘open (intr.)’
b. Hebrew	<i>picel</i> ‘split (tr.)’	<i>hit-pacel</i> ‘split (intr.)’
c. Swahili	<i>vunj-a</i> ‘break (tr.)’	<i>vunj-ik-a</i> ‘break (intr.)’

Anticausative constructions are similar to passive constructions in that the agent argument is “downgraded” (or “backgrounded”) (cf. Haspelmath 2022).

An example of a passive construction from Swahili

- (28) a. *Musa a-li-safisha nyumba.*
Musa 3SG.HUM-PST-clean house.G7
‘Musa cleaned the house.’
- b. *Nyumba i-li-safish-wa na Musa.*
house.G7 G7-PST-clean-PASS by Musa
‘The house was cleaned by Musa.’

Here the agent is expressed as an **oblique nominal**, but in an anticausative verb, the agent is completely absent.

Passive, *causative* and *anticausative* voice alternations can be schematized as follows (where antipassive is also shown):

<p>passive voice alternation:</p> $\begin{array}{ccc} \langle V & A & P \rangle \\ & & \\ \langle V\text{-PASS} & \{obl\} & S \rangle \end{array}$	<p>antipassive voice alternation:</p> $\begin{array}{ccc} \langle V & A & P \rangle \\ & & \\ \langle V\text{-ANTP} & S & \{obl\} \rangle \end{array}$
<p>causative voice alternation:</p> $\begin{array}{ccc} \langle V & & S \rangle \\ & & \\ \langle V\text{-CAUS} & A & P \rangle \end{array}$	<p>anticausative voice alternation:</p> $\begin{array}{ccc} \langle V & A & P \rangle \\ & & \\ \langle V\text{-ANTC} & & S \rangle \end{array}$

(from Haspelmath 2022)

(Note: This representation is not generally accepted in linguistics, and there is no universal agreement that transitivity should be discussed in terms of the role-types A (“agent”) and P (“patient”), rather than notions such as “subject” or “higher argument”; see Haspelmath 2011 for the latter.)

4. Labile verbs

Some languages use labile verbs, i.e. verbs (or verb pairs) that can be used in both senses, e.g. English

*We **opened** the window.*
*The window **opened**.*

Still more possibilities, e.g. suppletive pairs:

*We **killed** the bird.*
*The bird **died**.*

8. Extending the scale beyond change-of-state concepts

We have observed that some change-of-state concepts are more likely to be coded as causatives, and other are more likely to be coded as anticausatives (Haspelmath 1993):

'freeze'	>	'break'
more causative		more anticausative

This is because with 'freeze'-type verbs, the noncausal use is more frequent, while with 'break'-type verbs, the causal use is more frequent:

'Freeze' is "more spontaneous" than 'break'.

Now we can extend the scale to the left, because 'talk' is even more likely to be expressed by a causative:

'talk'	>	'freeze'	>	'break'
--------	---	----------	---	---------

In fact, it seems that in all languages (without exception), 'make someone talk' is expressed as a causative. (A few verbs of this sort, such as 'walk', are sometimes used as labile verbs, e.g. English *We walked the dog*.)

And we can extend the scale to the right, because 'be cut' is even more likely to be expressed as an anticausative (or "passive") – there are very few languages where there is an intransitive verb 'be cut':

'talk'	>	'freeze'	>	'break'	>	'be cut'
--------	---	----------	---	---------	---	----------

(the *spontaneity scale*, Haspelmath 2016)

In addition to presence of causative and anticausative coding, the scale corresponds to length of coding, because some languages have causatives of different lengths, e.g.

- (28) Halkomelem (Suttles 2004: 234-237)
 a. *méʔ* 'come off' *méʔ-x* 'take off'
 b. *ʔíməx* 'walk' *ʔíməx-stəx* 'make s.o. walk'
- (29) Japanese (Shibatani & Pardeshi 2002: §2.1)
 a. *kawak-* 'dry (intr.)' *kawak-as-* 'dry (tr.)'
 b. *hasir-* 'run' *hasir-ase-* 'make s.o. run'
- (30) Amharic: *a-* vs. *as-* (Amberber 2000: §3.1)
 a. *k'allət'ə* 'melt (intr.)' *a-k'allət'ə* 'melt (tr.)'
 b. *č'affərə* 'dance' *as-č'affərə* 'make s.o. dance'

(from Haspelmath 2016: 45)

The spontaneity scale thus makes several important predictions about the world's languages:

‘talk’	>	‘freeze’	>	‘break’	>	‘be cut’
		← more causative	more anticausative →			
		← longer causative	longer anticausative →			

9. Conclusion: Social-functional explanations of grammatical marking

I conclude that causative and anticausative verb formation are cross-linguistically quite variable, but there are certain tendencies, which can be explained by **efficiency of coding**:

Those verb meanings that are expressed **frequently** tend to be expressed by **shorter forms**, because they are more predictable, and speakers can afford to use less coding.

This is a **social-functional** theory, not a **cognitive-functional** theory. Most famous functional linguists (e.g. Givón 1984/1990; Tomasello 2003; Bybee 2010) take a cognitive-functional approach.

However, the **social and diachronic dimensions** of language are crucial for understanding language structures, too, as has been emphasized especially in Croft's (2000) more comprehensive functionalist approach.

References

- Alexiadou, Artemis, Elena Anagnostopoulou & Florian Schäfer. 2006. The properties of anticausatives crosslinguistically. In Mara Frascarelli (ed.), *Phases of interpretation*, 187–211. Berlin: Mouton de Gruyter.
- Bybee, Joan L. 2010. *Language, usage and cognition*. Cambridge: Cambridge University Press.
- Croft, William. 2000. *Explaining language change: An evolutionary approach*. Harlow: Longman.
- Givón, T. 1984/1990. *Syntax: A functional-typological introduction*. Amsterdam; Philadelphia: Benjamins
- Givón, T. 1991. Isomorphism in the grammatical code: Cognitive and biological considerations. *Studies in Language* 15(1). 85–114. <https://doi.org/10.1075/sl.15.1.04giv>.
- Haiman, John. 1983. Iconic and economic motivation. *Language* 59(4). 781–819. <https://doi.org/10.2307/413373>.
- Haspelmath, Martin. 1993. More on the typology of inchoative/causative verb alternations. In Bernard Comrie & Maria Polinsky (eds.), *Causatives and transitivity*, 87–120. Amsterdam: Benjamins.
- Haspelmath, Martin. 2011. On S, A, P, T, and R as comparative concepts for alignment typology. *Linguistic Typology* 15(3). 535–567.
- Haspelmath, Martin. 2016. Universals of causative and anticausative verb formation and the spontaneity scale. *Lingua Posnaniensis* 58(2). 33–63. <https://doi.org/10.1515/linpo-2016-0009>.
- Haspelmath, Martin. 2021. Explaining grammatical coding asymmetries: Form-frequency correspondences and predictability. *Journal of Linguistics* 57(3). 605–633. <https://doi.org/10.1017/S0022226720000535>.
- Haspelmath, Martin. 2022. Valency and voice constructions. Leipzig, MPI-EVA, ms. <https://ling.auf.net/lingbuzz/005941>.
- Jacobsen, Wesley M. 1985. Morphosyntactic transitivity and semantic markedness. *Chicago Linguistics Society* 21, Part 2. 89–104.
- Jacobsen, Wesley M. 2016. The semantic basis of Japanese transitive-intransitive derivational patterns. In Taro Kageyama & Wesley M. Jacobsen (eds.), *Transitivity and valency alternations: Studies on Japanese and beyond*, vol. 297, 445–478. Berlin: De Gruyter. <https://doi.org/10.1515/9783110477153-016>.
- Levshina, Natalia. 2023. *Communicative efficiency: Language structure and use*. Cambridge: Cambridge University Press.
- Matsumoto, Yo. 2016. Phonological and semantic subregularities in noncausative-causative verb pairs in Japanese. In Taro Kageyama & Wesley M. Jacobsen (eds.), *Transitivity and valency alternations: Studies on Japanese and beyond*. Berlin: De Gruyter. <https://doi.org/10.1515/9783110477153-016>.
- Mel'čuk, Igor. 1967. K ponjatiju slovoobrazovanija. *Izvestija AN SSSR: Serija literatury i jazyka* 26(4). 352–362.
- Nedjalkov, Vladimir P. & Georgij G. Sil'nickij. 1969. Tipologija morfoložičeskogo i leksičeskogo kauzativov [Typology of morphological and lexical causatives]. In Aleksandr A. Xolodovič (ed.), *Tipologija kauzativnyx konstrukcij [Typology of causative constructions]*, 20–60. Moskva: Nauka.
- Shibatani, Masayoshi (ed.). 1976. The grammar of causative constructions. In. New York: Academic Press.
- Shibatani, Masayoshi. 2016. The role of morphology in valency alternation phenomena. In Taro Kageyama & Wesley M. Jacobsen (eds.), *Transitivity and valency alternations: Studies on Japanese and beyond*, 445–478. Berlin: De Gruyter. <https://doi.org/10.1515/9783110477153-016>.
- Tomasello, Michael. 2003. *Constructing a language: A usage-based theory of language acquisition*. Cambridge, MA: Harvard University Press.

Voice constructions:

4B. Passives

MARTIN HASPELMATH

Max Planck Institute for Evolutionary Anthropology (Leipzig)

1. Passive voice constructions and passive-like constructions

Two examples of passive alternations:

(1) Japanese

a. *Naomi-ga Ken-o tatai-ta.*
 Naomi-NOM Ken-ACC hit-PST
 ‘Naomi hit Ken.’

b. *Ken-ga Naomi-ni tatak-are-ta.*
 Ken-NOM Naomi-OBL hit-PASS-PST
 ‘Ken was hit by Naomi.’ (passive construction)

(2) Turkish

a. *Musa ev-i temizle-di.*
 Musa house-ACC clean-PST
 ‘Musa cleaned the house.’

b. *Ev temizle-n-di.*
 house clean-PASS-PST
 ‘The house was cleaned.’ (passive construction)

Examples of passive-like alternations, lacking a passive marker on the verb (“uncoded”):

(3) Jamaican Creole (Kouwenberg 2023: 241, 244)

a. *Mieri rait di leta.*
 Mary write the letter
 ‘Mary wrote the letter.’

b. *Di leta rait.*
 the letter write
 ‘The letter has been written.’ (passive-like construction)

(4) Bambara (Cobbinah & Lüpke 2012: 136)

a. *Ù bε nɔ̃ dan.*
 3PL PRS millet sow
 ‘They sow millet.’

b. *Nɔ̃ bε dan (u fε).*
 millet PRS sow (they by)
 ‘Millet is sown (by them).’ (passive-like construction)

And, for comparison, an **antipassive** construction:

(5) Warronggo antipassive voice alternation (Tsunoda 2011: 428)

- | | | |
|-----------------------------|----------------|-----------------------|
| a. <i>Bama-nggo</i> | <i>gamo-Ø</i> | <i>yangga-n.</i> |
| man-ERG | water-ABS | search-NF |
| ‘The man looked for water.’ | | |
| b. <i>Bama-Ø</i> | <i>gamo-wu</i> | <i>yangga-gali-n.</i> |
| man-ABS | water-DAT | search-ANTP-NF |
| ‘The man looked for water.’ | | |

Question:

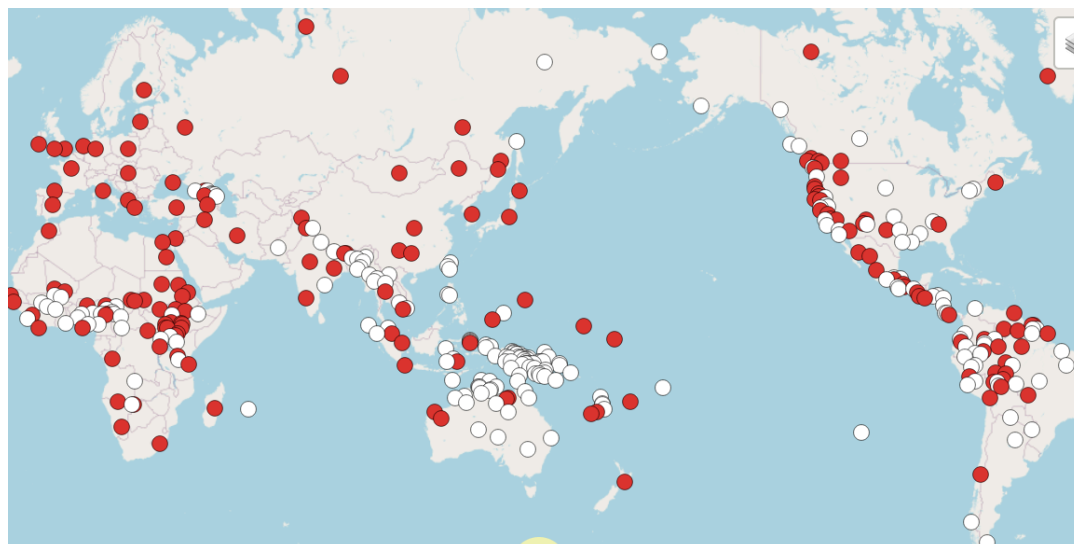
- How do we compare such constructions **across the world’s languages**?
- Can we say anything about **passive universals**?

3. Passives in the world’s languages

Passive constructions have been compared since von der Gabelentz’s (1861) typological study of passives.

But the modern systematic study of passives begins in 1975/1977 (Keenan 1975, Perlmutter & Postal 1977), when some American linguists argued that “universal grammar” should be based on “**grammatical relations**” rather than constituency structure. The first book on passives worldwide was Siewierska (1984).

Siewierska is also the author of the 2005 WALS map:



In Siewierska’s sample of 373 languages, 162 languages have a passive, while 211 lack a passive.

But what exactly is a passive construction?

Four proposed definitions:

A. Keenan (1975; 1985), in a modern version:

“Here I adopt the definition in Keenan (1985: 273) and define passive constructions as **subject-backgrounding operations** that derive n-place predicates from n+1 place predicates.” (Cabredo Hofherr 2023)

B. Comrie (1988: 12)

“I will **not attempt to set up necessary-and-sufficient conditions** to delimit [the passive], for the simple reason that I do not believe that such conditions are viable – while they can be established by fiat, there is no reason to suppose that any such definition will be linguistically significant.”

C. Haspelmath (1990: 27)

“A construction is called passive if (i) the **active subject** corresponds either to a **non-obligatory oblique phrase or to nothing**; and (ii) the active direct object (if any) **corresponds to the subject** of the passive; and (iii) the construction is somehow restricted vis-à-vis another unrestricted construction (the active), e.g. less frequent, functionally specialized, not fully productive.”

D. Siewierska (2005: 434)

“A construction has been classified as passive if it displays the following five properties:

- it contrasts with another construction, the active;
- the subject of the active corresponds to a **non-obligatory oblique phrase** of the passive **or is not overtly expressed**;
- the subject of the passive, if there is one, **corresponds to the direct object** of the active;
- the construction is pragmatically restricted relative to the active;
- the construction displays some **special morphological marking** of the verb.”

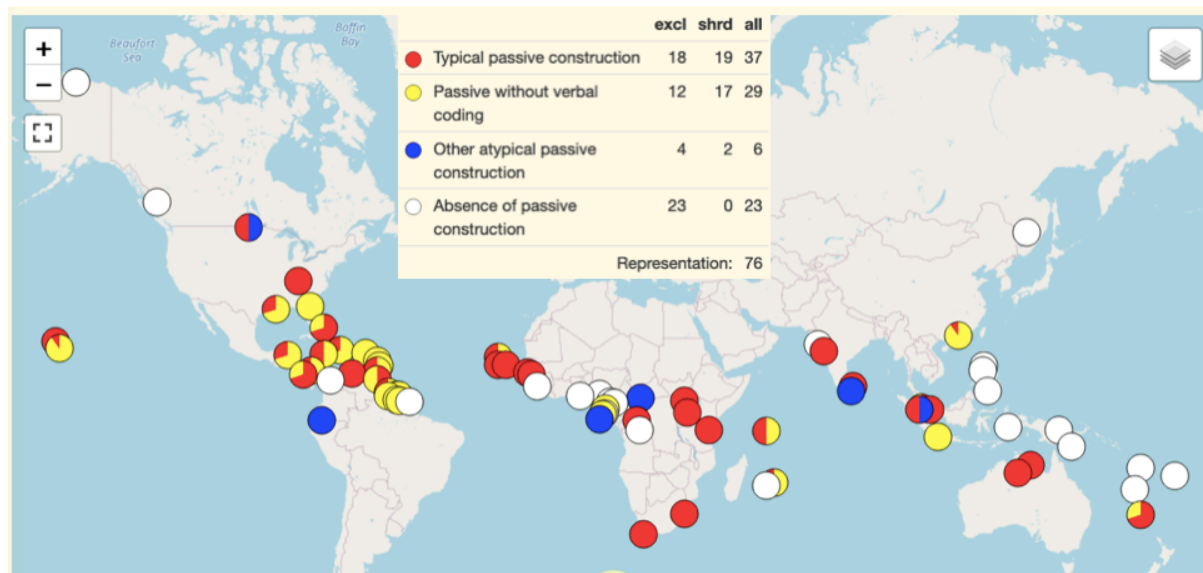
Similarly in **Grambank**: “Is there a morphological passive marked on the lexical verb?” (GB147, Skirgård et al. 2023)

Three questions:

- should the definition of “passive” make reference to “**operations**”?
- should it include the condition of **morphological marking**?
- should it be “**prototype**”-based (or perhaps “canonical”), **sharply** delineated?

Passives in the Atlas of Pidgin and Creole Language Structures (APiCS)

76 creole, pidgin and mixed languages (Michaelis et al. (eds) 2013); Haspelmath & APiCS Consortium (2013):



Feature value 1: **typical passive construction** (= by Siewierska's definition)

Feature value 2: **passive without verbal coding** (= oblique-agent construction)

In addition to “typical passive constructions”, quite a few Creole languages have uncoded passive-like constructions, e.g.

Principense (Portuguese-based)

Upanu me sa lava na abya.
Upanu me sa lava na abya.
 clothes POSS.1SG PROG wash LOC river
 ‘My clothes are being washed in the river.’

Virgin Islands Dutch Creole (19th century)

Ekeren fa roto sa mata di dungku da.
Ekeren fa roto sa mata di dungku da.
 each of rat FUT kill DET night there
 ‘Every rat would be killed that night.’

Haitian Creole (French-based)

Lakou a bale deja.
Lakou a bale deja.
 courtyard DEF sweep already
 ‘The court has already been swept.’
 French: ‘La cour est déjà balayée.’

4. Four possible passive universals

The following are some of the universals that have been proposed in the literature:

Universal 1

If a language has passives of **intransitive verbs** then it has passives of **transitive verbs**. (Keenan 1985)

Universal 2

In a passive construction, the verb always has a **passive marker**. (Haspelmath 1990).

Universal 3

If a passive marker can additionally have **reflexive** function, then it can also have **anticausative** function (e.g. Haspelmath 1987: 35).

Universal 4

If active and passive constructions differ with respect to topicality of agent and patient, then the patient is more topical in the passive, or the agent is less topical, or both. (cf. Haspelmath 2021: 155; Haspelmath 2025)

If we want to test these universals, we need

- a clear definition of “passive” with **sharp boundaries**
- a definition that can be **applied uniformly**, i.e. a *comparative concept*
 - this means that we cannot define passive in terms of an “**operation**” (because the evidence for transformations is language-particular)

This rules out the Keenan definition (which relies on an “**operation**”) and the Comrie approach (which is content with **unclear boundaries**).

5. *Passive* defined as a comparative concept for general linguistics

Following Haspelmath (2026), I now use the following definitions (which are actually very similar to Siewierska’s 2005):

(3) a. **passive voice alternation**

In a passive voice alternation, the uncoded alternant is transitive (with A and P), and in the coded alternant, the basic A is downgraded, and the basic P is promoted to S.

b. **passive voice construction**

A passive voice construction is the coded alternant in a passive voice alternation.

Voice alternations (= pairs of sister constructions with affixal coding on the verb) are best illustrated by vertical correspondence diagrams (Haspelmath 2026):

<p>passive voice alternation:</p> $\begin{array}{ccc} \langle V & A & P \rangle \\ & & \\ \langle V\text{-PASS} & \{\text{obl}\} & S \rangle \end{array}$	<p>antipassive voice alternation:</p> $\begin{array}{ccc} \langle V & A & P \rangle \\ & & \\ \langle V\text{-ANTP} & S & \{\text{obl}\} \rangle \end{array}$
<p>causative voice alternation:</p> $\begin{array}{ccc} \langle V & & S \rangle \\ & & \\ \langle V\text{-CAUS} & A & P \rangle \end{array}$	<p>anticausative voice alternation:</p> $\begin{array}{ccc} \langle V & A & P \rangle \\ & & \\ \langle V\text{-ANTC} & & S \rangle \end{array}$

But they can also be represented in a different, linear notation:

- (4) a. passive $\langle V, A_x, P_Y \rangle$
 $\approx \langle V\text{-PASS}, \{\text{obl}\}_x, S_Y \rangle$ ‘(X) acts on Y’
- b. antipassive $\langle V, A_x, P_Y \rangle$
 $\approx \langle V\text{-ANTP}, S_x, \{\text{obl}\}_Y \rangle$ ‘X acts (on Y)’
- c. causative $\langle V, S_x \rangle$ ‘X acts’
 $\approx \langle V\text{-CAUS}, A_z, P_x \rangle$ ‘Z makes X act’
- d. anticausative $\langle V, A_x, P_Y \rangle$ ‘X causes Y to change’
 $\approx \langle V\text{-ANTC}, S_Y \rangle$ ‘Y changes’

These definitions have the advantage that they can be applied to all languages using the same criteria:

- They do not make use of a notion of “subject”, which is notoriously non-universal, but instead use S, A and P (Haspelmath 2011).
- They do not make use of the notion of “inflection”, which cannot be defined clearly in such a way as to distinguish “inflectional voice categories” reliably from “derivational valency-changing constructions” (Haspelmath 2023).
- Note also that they do not make reference to “active voice”, but that they treat antipassive voice constructions, causative voice constructions and anticausative voice constructions in a parallel way.

A passive construction is a construction with a verbal affix and argument frame “S/oblique” that alternates with a non-affixed construction with argument frame “A/P”.

This is exactly parallel to *antipassive* (which always includes a verbal affix), as well as *causative* and *anticausative*.

Some people talk about “**uncoded antipassives**”:

The hunter shot the bear – *The hunter shot at the bear.*

or “**uncoded causatives**”:

The bear died. – *The hunter killed the bear.*

or “**uncoded anticausatives**”:

They broke the window. – *The window broke.*

However, all these terms originally referred to *constructions with special coding*, i.e. to *strategies*, not to functionally defined constructions (see Croft 2022; Haspelmath 2025 on *(construction-)strategies* versus *construction(-function)s*).

Thus, for terminological **transparency** (and **continuity** with the tradition), it is best not to extend the definition of “passive” to constructions where there is no coding on the verb.

6. What does this mean for the four passive universals?

A seemingly an unfortunate result of the decision to define “passive” in this way:

Universals 1 and 2 are no longer universals, because they are definitional:

Universal 1

If a language has passives of **intransitive verbs** then it has passives of **transitive verbs**. (Keenan 1985)

– passives **must** correspond to transitive verbs!

Universal 2

In a passive construction, the verb always has a **passive marker**. (Haspelmath 1990)

– passives **must** have a passive marker!

But “passives of intransitives” would not be called passives anyway, e.g. Russian

<i>on spit</i>	[he sleeps]	‘he sleeps’
<i>emu spit-sja</i>	[to.him sleeps-REFL]	‘he feels sleepy’

And Universal 2 is not true anyway, because **many uncoded passive-like constructions have been found in the meantime** (Holm 1997; Cobbinah & Lüpke 2012; Arkadiev 2024).

Instead of “uncoded passives”, I now talk about “passive-like constructions”.

7. Subjects versus topics

Another issue:

How do we distinguish between P-arguments that are simply **topics**, and P-arguments that are in **“subject” position** (= that have become S-arguments)?

Consider the “passive-like construction” in Lezgian (a Nakh-Dagestanian language of the Caucasus):

(11) Lezgian (Haspelmath 1991: 8)

a. *Farida ktab qaču-na.*
 Farid(ERG) book(ABS) take-AOR
 ‘Farid bought the book.’

b. *Ktab qaču-na.*
 book(ABS) take-AOR
 ‘The book was bought / One bought the book.’

Here we would not say that the P-argument has been shifted into S-argument position – a simpler view is that the A-argument has simply be omitted here (“withheld”, Haspelmath 2026).

But some authors have provided extensive argumentation, FOR or AGAINST subject status of the preverbal patient nominal in passive-like constructions.

FOR subject status: Donohue (2005) on Palu’e (Austronesian):

Unmarked clause type in Palu’e

A V P

(1) *Ia cube vavi vaʔa.*
 3SG shoot pig that
 ‘He shot that pig.’

Marked clause type in Palu’e

P A V

(2) *Vavi vaʔa ia cube.*
 pig that 3SG shoot
 ‘That pig, he shot (it).’ OR: ‘That pig was shot by him.’

The arguments are sometimes very subtle, e.g. involving a special type of quantifier floating (Donohue 2005: 71):

(59) *Keʔo(-keʔo) konen vaʔa ka-naba tetiʔón.*
 corn-RED 3PL that eat-all all
 ‘They all ate the corn.’
 * ‘They ate all of the corn.’

For Mauritian Creole, Syea (2024) argues AGAINST subject status, for example by noting that preverbal patient nominals cannot be indefinite:

- (25) a. loto la finn met dan garaz
car DEF PRF put in garage
'The car has been/was put in the garage.'
- b. lakaz la ti pe ankor ranze
house DEF PST PROG still build
'The house was still being built.'
- (26) a. *enn loto finn met dan garaz
a car PRF put in garage
'A car has been/was put in the garage.'
- b. *enn lakaz ti pe ankor ranze
a house PST PROG still build
'A house was still being built.'

I take this as another argument for restricting the well-established term *passive* to constructions with verbal marking.

8. Universals 3 and 4

Universals 3 and 4 are still testable, and quite possibly true generalizations:

Universal 3

If a passive marker can additionally have **reflexive** function, then it can also have **anticausative** function (e.g. Haspelmath 1987: 35).

Universal 4

If active and passive constructions differ with respect to topicality of agent and patient, then **the patient is more topical in the passive**, or **the agent is less topical**, or both. (cf. Haspelmath 2021: 155; Haspelmath 2025)

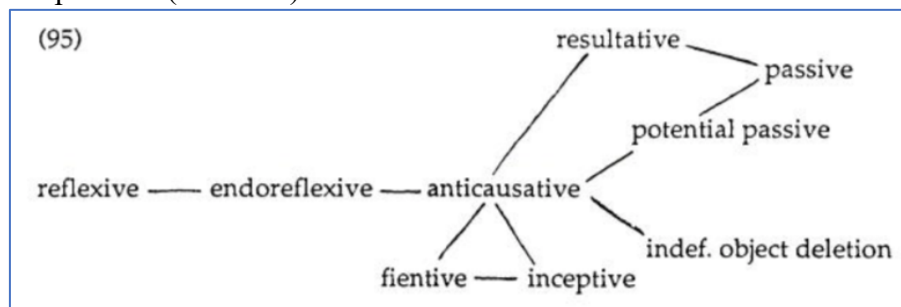
8.1. Universal 3: Reflexive–anticausative–passive coexpression (“polysemy”)

(15) Russian

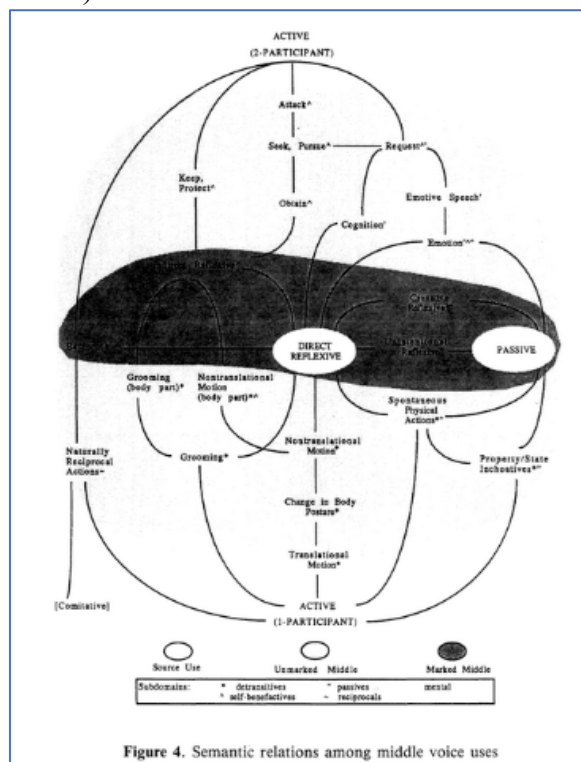
- a. *Mal'čik moet-sja.*
boy washes-REFL
'The boy washes (himself).'
- b. *Palka lomaet-sja.*
stick breaks-ANTC
'The stick breaks (intr).'
- c. *Dom stroit-sja.*
house builds-PASS
'The house is being built.'

Since Haspelmath (1987) and Croft et al. (1987), linguists have made **coexpression maps** about reflexive–passive “polysemy” (called “semantic maps”; Haspelmath 2003; Georgakopoulos & Polis 2018).

Haspelmath (1987: 35)



Croft et al. (1987: 187):



It may well be that the explanation for these coexpression maps lies in diachronic regularities (Cristofaro 2010) – reflexive markers always pass through an anticausative phase before they extend to passive uses (see also Bahrt 2021 on coexpression or “syncretism” patterns in voice constructions).

8.3. Universal 4: Deviations from usual associations between role rank and referential prominence

Usually, agents are more topical, and patients less topical. When this is reversed, languages tend to use extra coding, e.g. verb coding and oblique coding.

Universal 4 is actually a special case of the much more general Universal 5:

Universal 4

If active and passive constructions differ with respect to **topicality** of agent and patient, then the patient is more topical in the passive, or the agent is less topical, or both.

See Haspelmath (2021a) for the role-reference universal:

Universal 5: The role-reference association universal

Deviations from usual associations of role rank and referential prominence tend to be coded by longer grammatical forms if the coding is asymmetric.

Role rank:	A > P	(and R > T in ditransitives)
Referential prominence:	topical > nontopical	
	animate > inanimate	
	human > nonhuman	
	definite > indefinite	
	...	

Another prominent instantiation of this universal: *differential object flagging*

References

- APiCS: Michaelis, Susanne Maria, Philippe Maurer, Martin Haspelmath & Magnus Huber (eds.). (2013). *Atlas of pidgin and creole language structures*. Oxford: Oxford University Press.
- Bahrt, Nicklas N. 2021. *Voice syncretism*. Berlin: Language Science Press. (<https://langsci-press.org/catalog/view/315/3359/2373-1>)
- Cabredo Hofherr, Patricia Cabredo. 2023. Morphology of passives. *The Wiley Blackwell Companion to Morphology*. Wiley. (doi:10.1002/9781119693604.morphcom049)
- Cobbinah, Alexander & Friederike Lüpke 2012. Not cut to fit: Zero coded passives in African languages. In Matthias Brenzinger & Anne-Maria Fehn (eds.), *Proceedings of the 6th World Congress of African Linguistics*, 153–165. Cologne: Köppe.
- Comrie, Bernard. 1988. Passive and voice. In Shibatani, Masayoshi (ed.), *Passive and voice*, vol. 1, 9–23. Amsterdam: Benjamins.
- Creissels, Denis. 2024. *Transitivity, valency and voice*. Oxford: Oxford University Press (to appear).
- Croft, William. 2022. *Morphosyntax: Constructions of the world's languages*. Cambridge: Cambridge University Press.
- Donohue, Mark. 2005. The Palu'e passive: From pragmatic construction to grammatical device. In Arka, I. Wayan & Ross, Malcolm (eds.), *The many faces of Austronesian voice systems: Some new empirical studies*, 59–86. Canberra: Australian National University.
- Haspelmath, Martin. 1987. *Transitivity alternations of the anticausative type* (Arbeitspapiere Des Instituts Für Sprachwissenschaft N.F. Nr. 4). Cologne: Universität zu Köln. (<http://doi.org/10.5281/zenodo.225196>)
- Haspelmath, Martin. 1990. The grammaticization of passive morphology. *Studies in Language* 14(1). 25–72.
- Haspelmath, Martin. 1991. On the question of deep ergativity: The evidence from Lezgian. *Papier zur Linguistik* 44/45(1–2). 5–27. (doi:10.5281/zenodo.225289)
- Haspelmath, Martin. 2021. Role-reference associations and the explanation of argument coding splits. *Linguistics* 59(1). 123–174. (doi:10.1515/ling-2020-0252)
- Haspelmath, Martin. 2022. Valency and voice constructions. Leipzig, MPI-EVA, ms. <https://ling.auf.net/lingbuzz/005941>.
- Haspelmath, Martin. 2024. Construction-functions and construction-strategies. In Däbritz, Chris Lasse & Budzisch, Josefine & Basile, Rodolfo (eds.), *Existential and locative constructions*. Berlin: Language Science Press (to appear). (<https://ling.auf.net/lingbuzz/007897>)
- Haspelmath, Martin. 2025. Passive constructions: Universals and language description. To appear. (<https://ling.auf.net/lingbuzz/008340>)
- Holm, John. 1997. Passive-like constructions in English-based and other creoles. In Edgar Werner Schneider (ed.), *Englishes around the world (Volume 1: General studies, British Isles, North America)*. Amsterdam: Benjamins.

- Keenan, Edward L. 1975. Some universals of passive in Relational Grammar. In Grossman, R. & San, L.J. & Vance, T. (eds.), *Proceedings of the Eleventh Regional Meeting of the Chicago Linguistic Society*, 340–352. Chicago: Chicago Linguistic Society.
- Keenan, Edward L. 1985. Passive in the world's languages. In Shopen, Timothy (ed.), *Language typology and syntactic description (Volume I)*, 243–281. Cambridge: Cambridge University Press.
- Kouwenberg, Silvia. 2023. The creole passive. In Patricia Cabredo Hofherr, Herby Glaude & Elena Soare (eds.), *Grammar is a moveable feast: A webschrift for Anne Zribi-Hertz*, 239–249. Paris: Zenodo. <https://doi.org/10.5281/zenodo.8055461>.
- Kulikov, Leonid I. 2011. Voice typology. In Song, Jae Jung (ed.), *The Oxford handbook of linguistic typology*, 368–398. Oxford: Oxford University Press.
- Mel'čuk, Igor A. & Xolodovič, Aleksandr A. 1970. K teorii grammatičeskogo zaloga. *Narody Azii i Afriki* 1970(4). 111–124.
- Michaelis, Susanne Maria & Maurer, Philippe & Haspelmath, Martin & Huber, Magnus (eds.). 2013. *Atlas of pidgin and creole language structures*. Oxford: Oxford University Press.
- Perlmutter, David M. & Postal, Paul M. 1977. Toward a universal characterization of passivization. *Annual Meeting of the Berkeley Linguistics Society*, vol. 3, 394–417.
- Siewierska, Anna. 1984. *The passive: A contrastive linguistic analysis*. London: Croom Helm.
- Siewierska, Anna. 2005/2013. Passive constructions. In Matthew S. Dryer & Martin Haspelmath (eds.), *The world atlas of language structures online*. Leipzig: MPI-EVA. <https://wals.info/chapter/107>.
- Skirgård, Hedvig & Haynie, Hannah J. & Blasi, Damián E. & Hammarström, Harald & Collins, Jeremy & Latache, Jay J. & Lesage, Jakob et al. 2023. Grambank reveals the importance of genealogical constraints on linguistic diversity and highlights the impact of language loss. *Science Advances* 9(16). eadg6175. (doi:10.1126/sciadv.adg6175)
- Syea, Anand. 2024. Absence of syntactic passive in creoles: Evidence from French-based Mauritian Creole. *Canadian Journal of Linguistics/Revue canadienne de linguistique* 1–29. (doi:10.1017/cnj.2024.13)
- von der Gabelentz, Hans Conon. 1861. Über das Passivum: Eine sprachvergleichende Abhandlung. *Abhandlungen der Philologisch-historischen Classe der Königlich sächsischen Gesellschaft der Wissenschaften* 3. 449–546.
- Winford, Donald. 1993. *Predication in Caribbean English creoles*. Amsterdam: John Benjamins.
- Zúñiga, Fernando & Seppo Kittilä. 2019. *Grammatical voice*. Cambridge: Cambridge University Press.

Voice constructions:

4C. Antipassives and inverses

MARTIN HASPELMATH

Max Planck Institute for Evolutionary Anthropology (Leipzig)

1. Examples of antipassives

(1) Warronggo antipassive voice alternation (Tsunoda 2011: 428)

a. *Bama-nggo gamo-Ø yangga-n.*

man-ERG water-ABS search-NF

‘The man looked for water.’

b. *Bama-Ø gamo-wu yangga-gali-n.*

man-ABS water-DAT search-ANTP-NF

‘The man looked for water.’ (antipassive construction)

(2) K’iche’ (Mayan; Mondloch 1981; cited after Heaton 2020: 140)

Transitive

K-Ø-in-yoq’	le:	in-ta:t.
INCOMPL-3 SG.ABS-1 SG.ERG-mock	DET	1 SG.POSS-father
‘I mock my father.’		

Pragmatic use of the antipassive

K-in-yoq’-on	(č-e:	le:	in-ta:t).
INCOMPL-3 SG.ABS-mock- <u>ANTIP</u>	<u>OBL</u> -3 SG	DET	1 SG.POSS-father
‘I mock (my father).’			
(Mondloch 1981, p. 171)			

Syntactic use of the antipassive

Xač’in	š-Ø-yoq’-on	č-e:	ri:	išoq’
WH	COMPL-3 SG.ABS-mock- <u>ANTIP</u>	<u>OBL</u> -3 SG	DET	woman
‘Who mocked the woman?’				
(Mondloch 1981, p. 189)				

2. A definition of *antipassive construction* as a strategy

Since the 1970s, antipassive voice constructions have been defined in a way that is analogous to passive constructions: an antipassive is **the mirror image** of a passive, so the definition is ideally parallel:

(3) a. **antipassive voice alternation**

In an antipassive voice alternation, the uncoded alternant is transitive (with A- and P-arguments), where in the coded alternant, the basic P is downgraded, and the basic A corresponds to S.

b. **antipassive voice construction**

An antipassive voice construction is the coded alternant in an antipassive voice alternation.

Thus, the notion of an **antipassive voice construction** presupposes the notion of an **antipassive voice alternation**.

Some antipassives are **P-demoting** (with the argument corresponding to the basic P marked by an oblique flag):

(4) Nyawaygi (Pama-Nyungan; Dixon 1983: 496)

a. *ŋin-da yagu bu:ga*
 you-ERG water.ABS drink
 ‘You drink some water!’

b. *ŋin-ba bu:g-iga yagu-gu*
 you-ABS drink-ANTP water-DAT
 ‘You drink some water!’

... while others are **P-suppressing** (with the basic P eliminated from the valency):

(5) Russian (Say 2005: 266)

a. *Ne dav-i ètu pačku.*
 not squeeze-IMPV this pack
 ‘Don’t squeeze this pack.’

b. *Ne dav-i-s’.*
 not squeeze-IMPV-ANTP
 ‘Don’t squeeze.’

The distinction between P-demoting and P-suppressing antipassives paralleled in passive constructions, which may be A-demoting (as in English, where an oblique agent phrase is allowed), or A-suppressing.

The definition in (3) specifies that the antipassive construction is **intransitive** (with an S-argument), while the uncoded alternant is **transitive** (with an A-argument and a P-argument).

(See, e.g. Dixon 1994: 146; Polinsky 2017: 309; Janic & Witzlack-Makarevich 2021: 2; Seržant et al. 2021: 970; Creissels 2024: 428),

Some authors work with a notion of “semantic transitivity” (e.g. Polinsky 2017: 310; Janic & Witzlack-Makarevich 2021: 3; Creissels 2024: §3.1.1), but it is best to use the terms *transitive*, *passive* and *antipassive* for kinds of **(formally defined) strategies**, and not for the functions (or **functionally defined constructions**) that they express.

A voice alternation such as antipassive is a pair of two different valency patterns, an **uncoded** and a **coded** one (these are two strategies for expressing roughly the same content).

- (6) antipassive $\langle V, A_X, P_Y \rangle$
 $\approx \langle V\text{-ANTP}, S_X, \{\text{obl}\}_Y \rangle$ ‘X acts (on Y)’
- (7) passive $\langle V, A_X, P_Y \rangle$
 $\approx \langle V\text{-PASS}, \{\text{obl}\}_X, S_Y \rangle$ ‘(X) acts on Y’

The **constructional view** presented here is neutral between a “syntactic” and a “lexical(ist)” view that have sometimes been contrasted (e.g. Polinsky 2017: §13.4).

Constructions have the same properties in “**syntax**”, “**morphology**”, and **lexicon**, and the “lexicon-grammar” contrast is not well-motivated (Haspelmath 2023: §6).

• **pivot-creating role of antipassives**: in some languages an antipassive construction must be used for relativizing or questioning the agent or other related purposes.

E.g. in Inuktitut (Johns 1992), the antipassive construction is required for relativizing the agent argument:

- (8) Inuktitut (Johns 1992: 61, 72)
- a. *Anguti-up nanuq kapi-ja-a.*
 man-ERG bear.ABS stab-PTCP-3SG>3SG
 ‘The man stabbed the bear.’
- b. **Anguti(-up) nanuq kapi-ja-a ani-juq.*
 man(-ERG) bear.ABS stab-PTCP-3SG leave-PTCP.3SG
 (‘The man who stabbed the bear left.’)
- c. *Anguti nanur-mik kapi-si-juq ani-juq.*
 man.ABS bear-MOD stab-ANTP-PTCP.3SG leave-PTCP.3SG
 ‘The man who stabbed the bear left.’

3. Three potential antipassive universals

We can now consider three potential universals of antipassive constructions. All three are implicational universals (“preferentially” is intended as an abbreviatory device).

- (9) Universal 1: **Antipassive non-specificity**
 P-demoting antipassives preferentially occur with indefinite and/or non-specific patient arguments.
- (10) Universal 2: **Antipassive atelicity**
 Antipassives preferentially occur with atelic and/or imperfective aspectual reference.
- (11) Universal 3: **Secondary pivot-creating antipassives**
 If a language has an antipassive construction that is required for some pivot constructions, the antipassive also occurs outside of these pivot constructions.

These universals are well-known from the literature on antipassives (e.g. Cooreman 1994; Polinsky 2013; Heaton 2020; Janic & Witzlack-Makarevich 2021), though they have not often been formulated in this way.

Most general works on antipassives make **general observations**, such as the following:

“The main functions [of antipassives] discussed in the literature are related to semantic properties such as the indefiniteness or low identifiability of the object and imperfective aspect of the predicate, or to syntactic functions such as the realignment of the subject from transitive to intransitive, and topic chaining.” (Murasugi 2020: 277)

Here I stress the advantages of formulations in terms of **universals**, as they potentially tell us something about Human Language in general, and they are (at least in principle) falsifiable hypotheses.

For **Universal 1 (Antipassive non-specificity)**, one can cite:

(12)a. Murasugi (2020: 277)

“One of the most common correlates with the use of the antipassive is the low degree of identifiability of the object (Cooreman 1994). Objects with low identifiability have been characterized in a number of ways, including indefinite, non-referential, nonspecific, generic, unknown, and undisclosed; essentially, their identity is not considered an important part of the discourse.”

For **Universal 2 (Antipassive atelicity)**, one can cite the following:

(13)a. Creissels (2024: 434)

“In many languages, antipassivization implies reference to habitual rather than specific events, and the expression of the participant that would be encoded as P in the transitive construction, if possible at all, is conditioned by a low degree of referentiality. These tendencies have been repeatedly pointed out in the literature on antipassive.

Universal 2 is formulated by Polinsky (2017) as an implicational universal (something that is done rarely in the literature on antipassives):

(14) “If an antipassive construction can have a perfective (telic) interpretation, it must also have an imperfective (non-telic) interpretation.” (Polinsky 2017: 316)

Universal 3 about pivot-creating antipassives:

(15) a. Janic & Witzlack-Makarevich (2021: 23)

“if a language has the syntactically motivated antipassive construction, it seems also to use the antipassive construction for semantic and/or discourse-pragmatic functions. ... The reverse situation is not necessarily true: the presence of a semantically and discourse pragmatically motivated antipassive construction does not imply that a language will also have the antipassive construction used for the syntactic reasons.”

Pivot-creating antipassives were prominent in the early phases of research on antipassives (on languages such as Dyirbal, Chukchi, varieties of Inuktut-Yupik, varieties of Mayan), but it seems that they are **far rarer than other kinds of antipassives** (Heaton 2020: §3.3).

4. Antipassive as the mirror-image of passive

the definition given here presents antipassive constructions as an exact mirror-image of passive constructions, cf.

(16)a. **passive voice alternation**

In a passive voice alternation, the uncoded alternant is transitive (with A- and P-arguments), while in the coded alternant, the basic A is downgraded, and the basic P corresponds to S.

b. **passive voice construction**

A passive voice construction is the coded alternant in a passive voice alternation.

- these definitions do not make use of a notion of “subject”, in contrast with, e.g., Polinsky (2026):

“The passive canonically “demotes” the **logical subject (agent)**, either removing it entirely or expressing it as an oblique phrase (as in English “by-phrases”), while promoting the logical object (patient, undergoer) to **subject position**. In contrast, the antipassive is understood as an operation that demotes or omits the logical object while maintaining the prominence of the logical subject.” (Polinsky 2026)

BUT: While the oblique passive agent seems to be always **optional**, the demoted argument in antipassives is sometimes obligatory, e.g.

(17) French (Janic 2016: 205)

- Les étudiants attaquent les révisions.*
the students start the revisions
‘The students start the revisions.’
- Les étudiants s’attaquent aux révisions.*
the students ANTP start to.the revisions
‘The students start (with) the revisions.’

5. More on the definition of *antipassive construction*

Some authors offer **prototype-based definitions**, e.g. Zúñiga & Kittilä (2019: 103), as well as Polinsky (2026):

“In its **prototypical** form, the antipassive involves morphological marking on the verb, often through dedicated antipassive morphology, while the erstwhile patient (i.e., affected object) either becomes an oblique argument (e.g., a prepositional phrase) or is completely omitted from the clause.” (Polinsky 2026)

But non-sharp definitions like this one do not allow us to formulate **testable hypotheses**, and prototype-based definitions can be extended in an almost limitless way, thus leading to a very blurred picture.

Vigus (2018: 343) discusses “structural criteria” (i.e. criteria for defining antipassive as a **strategy** rather than as a functionally defined construction) and correctly observes:

“It is not clear why a particular combination of structural criteria should be considered as the only “true” type of antipassive. There also appears to be no principled way of deciding which criteria to select as definitional.”

Choosing a definition of a comparative-concept term is a **methodological decision**, and the choice need not be “principled”. Apart from sharpness and uniform applicability, a requirement for a definition is that it should be **clear** and **not too long**, and it should be easy to remember.

The definition given here is **easy to remember** also because it is parallel to the definition of *passive*, thus making the term *anti-passive* quite transparent.

- there are, of course, many antipassive-like constructions that one would like to bring into in the broader picture, including the English **conative alternation**:

shoot the bear/shoot at the bear

but a definition must primarily be clear and sharp. For a broader picture, one could use a different term, such as “P-backgrounding construction” (e.g. Seržant et al. 2021).

6. Describing antipassive(-like) constructions: In the language’s own terms

Different languages have **different structural systems**, for example different syllable structures (e.g. Blevins 1995), different case systems (e.g. Blake 2001), and so on – this was emphasized by the classics of structuralism, e.g.

“None of [the concepts of grammar], as used till now, is susceptible of general definition: genitive, perfect, subjunctive, and passive are quite different things in one language, Latin for example, from what they are in another, say Greek. The same is true, without any exception, of the remaining concepts of conventional linguistics.” (Hjelmslev 1961: 12)

If each language is described with its own grammatical concepts, this means that **language descriptors are not dependent on general linguists for their analyses**. If they find a phenomenon that looks unfamiliar, they can give it a completely new name.

(23) Nez Perce (Sahaptian; Rude 1986: 126; 129)

a. *Háama-nm péé-'wi-ye' wewúkiye-ne.*
 man-ERG 3.TR-shoot-ASP elk-ACC
 'The man shot the elk.'

b. *Páaxloo ha'áyat hi-x'ni-siix qé'mes.*
 five women 3.NOM-dig-ASP camas
 'Five women are digging camas (*Camassia quamash*).'

Rude (1986: 129) calls the second construction “antipassive”, while Deal (2010: 75) initially calls them *caseless clauses*, though she also uses the term “antipassive”. But she notes:

“The Nez Perce antipassive differs from a classical antipassive construction in not showing oblique case on the object (which I take to be a low-level effect) and in not showing an antipassive morpheme in the verb.”
 (Deal 2010: 84)

But what is the point of using the term *antipassive* here if the construction does not fall under the definition of antipassive?

7. Conclusion on antipassives

The broad domain of **patient-backgrounding constructions** is heterogeneous. Heaton (2020 describes the findings of a study of over four hundred languages (Heaton 2017):

“The results were striking in terms of just how much variation there is in the data: Even though there were more than 100 different sets of features that describe different antipassive-like constructions, only 64 of these sets of features appeared twice or more in the data set. Therefore, many sets of features which describe a structure in a given language did not recur, highlighting the heterogeneity of this pool of constructions potentially lumped together under the antipassive label.”
 (Heaton 2020: 137-138)

Because of this heterogeneity, it is **quite likely that the prominence of the term *antipassive* is due to an accident of the history of linguistics**, not due to a special importance of the constructions that fall under this label.

More generally, there is no reason to think that one terminological system is necessarily better than others, but also **no reason to think that terminological confusions will become fewer automatically** as research on different languages advances.

8. Inverse constructions

8.1. Examples of inverse contrasts

(3) Itonama (Bolivia; Crevels 2023: 518)

a. non-inverse

E'-kamo'-ke Ihwana.

2SG-hit.face-PL Juan

'You hit Juan (several times in the face).' (2→3)

b. inverse

A'-k'i-kamo'-ke Ihwana.

2SG-INV-hit.face-PL Juan

'Juan hit you (several times in the face).' (3→2)

(4) Southern Tiwa (Tanoan; New Mexico; Rosen 1990: 686; 697)

a. non-inverse

Seuanide ti-mũ-ban.

man 1SG-see-PST

'I saw the man.' (1→3)

b. inverse

Seuanide-ba ti-mũ-che-ban.

man-INS 1SG-see-INV-PST

'The man saw me = I was seen by the man.' (3→1)

Proposed definition:

(2) **inverse** (inverse indexing system)

A language has an inverse indexing system if its set of agent-patient person scenario forms include a special marker in some or all upstream scenarios that does not occur in any of the downstream scenarios. (Haspelmath 2025)

8.2. Downstream and upstream scenarios

(4) balanced scenarios (both arguments with equal prominence): 3→3
 downstream scenarios (agent outranks patient): 1→2, 1→3, 2→3
 upstream scenarios (patient outranks agent): 2→1, 3→1, 3→2

	1ST PATIENT	2ND PATIENT	3RD PATIENT
1ST AGENT	—	<i>downstream</i>	<i>downstream</i>
2ND AGENT	<i>upstream</i>	—	<i>downstream</i>
3RD AGENT	<i>upstream</i>	<i>upstream</i>	<i>balanced</i>

Figure 1: Downstream, upstream and balanced person scenarios

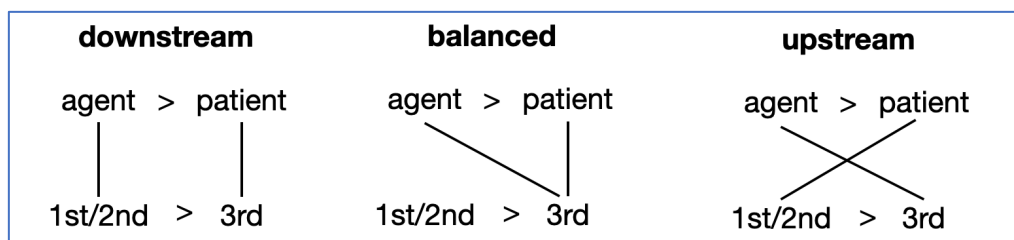


Figure 2: The association between agent/patient roles and the person scale

	LOCUPHORIC PATIENT (1ST/2ND)	3RD PATIENT
LOCUPHORIC AGENT (1ST/2ND)	(locuphoric domain)	mixed domain (<i>downstream</i>)
3RD AGENT	mixed domain (<i>upstream</i>)	aliophoric domain (<i>balanced</i>)

Figure 3: Locuphoric, mixed and aliophoric domains, with downstream, upstream and balanced person scenarios

8.3. On the “stream” and “flow” metaphor

“...the inverse is understood in terms of the relationship between the canonical event participants of a transitive clause: the agent and the patient, and is defined by event flow directionality vis-à-vis these participants. The inverse is taken to signal a reversion from the “normal” event flow with respect to a participant hierarchy, most commonly 1>2>3.” (Velázquez-Castillo 2008: 383)

The notion of “flow” seems to originate in DeLancey (1981), though he talked about “attention flow”.

The *upstream/downstream* terminology makes reference not only to a “flow” metaphor, but also to an “**action direction**” metaphor, as formulated by Haude & Zúñiga:

“When the direction of the action goes from a participant that ranks higher in [a prominence] hierarchy towards a lower-ranking participant, the direct construction is chosen (with an unmarked or direct-marked verb); when the action goes in the opposite direction, the inverse construction is chosen (with inverse marking on the verb.” (Haude & Zúñiga 2016: 444)

8.4. Is there an inverse voice?

Some linguists have said that there is a “category of direction” (direct/inverse), just as there is a “category of voice” (active/passive).

But what is a “category”? And what is “inflection”? (cf. Haspelmath 2024b)

Proposal for what a *voice alternation* is (Haspelmath 2022):

(5) voice alternation

A voice alternation is a valency alternation with coding of the valency by an affix on a verb.

This is a non-traditional meaning of the term *voice*, but it is now widely shared in syntactic typology (e.g. Kulikov 2011; Zúñiga & Kittilä 2019; Bahrt 2021; Creissels 2024).

(Note that Shibatani (2006) and Croft (2022) use *voice* in a rather different way.)

Notation used in Haspelmath (2022):

- | | | | |
|-----|--------------|---|---------------------------------------|
| (6) | labile | $\langle V, A_X, P_Y \rangle$
$\approx \langle V, S_Y \rangle$ | ‘X causes Y to change’
‘Y changes’ |
| (7) | a. passive | $\langle V, A_X, P_Y \rangle$
$\approx \langle V\text{-PASS}, \{\text{obl}\}_X, S_Y \rangle$ | ‘(X) acts on Y’ |
| | b. causative | $\langle V, S_X \rangle$
$\approx \langle V\text{-CAUS}, A_Z, P_X \rangle$ | ‘X acts’
‘Z makes X act’ |

According to this way of using the term *voice*, inverse constructions do not fall under the definition of voice, even though one could suggest something like (8):

- | | | | |
|-----|----------------|--|-----------------------|
| (8) | direct-inverse | $\langle V\text{-DIR}, A_X, P_Y \rangle$
$\approx \langle V\text{-INV}, A_Y, P_X \rangle$ | (direct)
(inverse) |
|-----|----------------|--|-----------------------|

Farrell (2005: 74) suggests such a view of inverse systems, but linguists usually make a distinction between alternations and coding splits:

- | | | |
|-----|-------------------------|--|
| (9) | a. alternation: | a set of two closely related constructions that can express roughly the same meaning and between speakers can choose |
| | b. coding split: | a set of different coding patterns that occur under different conditions |

On this view, inverse systems are systems with *split coding* (see also Haspelmath 2021), not alternations.

8.5. Against fuzzy (or prototypical) definitions

Jacques & Antonov (2014): a canonical inverse has the following properties:

- person indexes are role-neutral
- the conominals keep the same flagging
- the resulting ambiguity is resolved by inverse (or direct) markers

The first two are counterexemplified by Itonama (3) and Southern Tiwa (4b), and there is not necessarily any ambiguity.

It seems that Jacques & Antonov started out from a “prototypical” inverse as described for Algonquian languages, because their definition looks like a description of Cree.

However, why should Cree be a “prototype” for our thinking about kinds of grammatical marking in the world’s languages?

They do not say where the boundaries of the inverse are, and they even mention a “highly non-canonical inverse system” in Khaling (a Kiranti language) which deviates so much from the canon that they say that “it would be misleading to refer to it as an inverse” (2014: 308).

8.6. The inverse system universal

The chief purpose of comparative grammar is to identify cross-linguistically general properties, so here is a proposal:

(10) The inverse system universal

Anti-inverse systems do not exist.

This has often been implicit, but here it is perhaps fully explicit for the first time. Implicit here:

“When the direction of the action goes from a participant that ranks higher in [a prominence] hierarchy towards a lower-ranking participant, the direct construction is chosen (with an **unmarked or direct-marked verb**); when the action goes in the opposite direction, the inverse construction is chosen (with **inverse marking** on the verb.” (Haude & Zúñiga 2016: 444)

The inverse system universal is a special case of the highly general role-reference association universal:

(41) The role-reference association universal (Haspelmath 2021)

Deviations from usual associations of role rank and referential prominence tend to be coded by longer grammatical forms if the coding is asymmetric.

Compare also inverse-like patterns such as the obligatory passives described by Shibatani (2003: 278):

- (43) a. *Boku-wa Ken-o nagut-ta.* (Active)
 I-TOP Ken-ACC hit-PAST
 ‘I hit Ken.’
 b. **Ken-ga boku-o nagut-ta.* (Active)
 Ken-NOM I-ACC hit-PAST
 ‘Ken hit me.’
 c. *Boku-wa Ken-ni nagura-re-ta.* (Passive)
 I-TOP Ken-by hit-PASS-PAST
 ‘I was hit by Ken.’

This is not an inverse indexing system, because Japanese has no person indexing, but it is of course similar.

The similarities can be described in terms of the “downstream/upstream” terminology, and we do not need to extend the term *inverse*.

Note also: The Japanese example (43a) is completely unambiguous, but Japanese still requires the passive.

This is a more general feature of coding splits motivated by the role-reference association universal: Extra coding is often motivated by unusualness (rarity of occurrence), not by the goal of avoiding ambiguity.

e.g. Portuguese DOM	<i>eu</i>	‘I’
	<i>a mim</i>	‘me’

References

- Aldridge, Edith. 2012. Antipassive and ergativity in Tagalog. *Lingua* 122(3). 192–203.
- Anderson, Stephen R. 2016. Synchronic versus diachronic explanation and the nature of the Language Faculty. *Annual Review of Linguistics* 2(1). (doi:10.1146/annurev-linguistics-011415-040735)
- Bahrt, Nicklas N. 2021. *Voice syncretism*. Berlin: Language Science Press. <https://langsci-press.org/catalog/view/315/3359/2373-1>.
- Bickel, Balthasar & Fernando Zúñiga. 2017. The “word” in polysynthetic languages: Phonological and syntactic challenges. In *Handbook of polysynthesis*.
- Bickel, Balthasar. 2011. Grammatical relations typology. In Song, Jae Jung (ed.), *The Oxford handbook of linguistic typology*, 399–444. Oxford: Oxford University Press.
- Blake, Barry J. 1990. *Relational grammar*. London: Routledge.
- Blake, Barry J. 2001. *Case* (Cambridge Textbooks in Linguistics). Second edition. Cambridge: Cambridge University Press.
- Blevins, Juliette. 1995. The syllable in phonological theory. In Goldsmith, John (ed.), *The handbook of phonological theory*, 206–244. Cambridge, MA: Blackwell.
- Boas, Franz. 1911. Introduction. In Boas, Franz (ed.), *Handbook of American Indian languages (Volume 1)* (Bureau of American Ethnology Bulletin 40–1), 5–83. Washington: Government Printing Office.
- Carrier, Julien. 2017. The ergative-antipassive alternation in Inuktitut: Analyzed in a case of new-dialect formation. *Canadian Journal of Linguistics/Revue canadienne de linguistique* 62(4). 661–684. (doi:10.1017/cnj.2017.33)
- Chen, Victoria. 2017. *A reexamination of the Philippine-type voice system and its implications for Austronesian primary-level subgrouping*. Manoa, HI: University of Hawai‘i at Manoa. (PhD dissertation.)
- Chung, Sandra. 2025. Antipassive in a Minimalist universal grammar. In Clemens, Lauren & Gribanova, Vera & Scontras, Gregory (eds.), *Syntax in uncharted territories: Essays in Honor of Maria Polinsky*. Irvine: UC Irvine. (<https://escholarship.org/uc/item/2n4661zq>)
- Comrie, Bernard & Forker, Diana & Khalilova, Zaira & van den Berg, Helma. 2021. Antipassives in Nakh-Daghestanian languages: Exploring the margins of a construction. In Janic, Katarzyna & Witzlack-Makarevich, Alena (eds.), *Antipassive: Typology, diachrony, and related constructions*, 515–548. Amsterdam: Benjamins.
- Comrie, Bernard. 1981. Aspect and voice: Some reflections on perfect and passive. In Tedeschi, Philip J & Zaenen, Annie (eds.), *Tense and aspect* (Syntax and Semantics 14), 65–78. New York: Academic Press.
- Cooreman, Ann. 1994. A functional typology of antipassives. In Fox, Barbara A. & Hopper, Paul J. (eds.), *Voice: form and function*, 49–88. Amsterdam: Benjamins.

- Creissels, Denis. 2021. Antipassive derivation in Soninke (West Mande). In Janic, Katarzyna & Witzlack-Makarevich, Alena (eds.), *Antipassive: Typology, diachrony, and related constructions*, 293–314. Amsterdam: Benjamins.
- Creissels, Denis. 2024. *Transitivity, valency and voice*. Oxford: Oxford University Press.
- Creissels, Denis. 2024. *Transitivity, valency and voice*. Oxford: Oxford University Press.
- Crevels, Mily. 2023. Itonama. In Patience Epps & Lev Michael (eds.), *Amazonian Languages: Language isolates I (Aikanã to Kandozi-Shapra)*. Berlin: De Gruyter. <https://doi.org/10.1515/9783110419405>.
- Croft, William. 2001. *Radical Construction Grammar: Syntactic theory in typological perspective*. Oxford: Oxford University Press.
- Croft, William. 2009. Methods for finding universals in syntax. In Scalise, Sergio & Magni, Elisabetta & Bisetto, Antonietta (eds.), *Universals of language today*, 145–164. Dordrecht: Springer.
- Croft, William. 2016. Comparative concepts and language-specific categories: Theory and practice. *Linguistic Typology* 20(2). 377–393. (doi:10.1515/lingty-2016-0012)
- Croft, William. 2022. *Morphosyntax: Constructions of the world's languages*. Cambridge: Cambridge University Press.
- Deal, Amy Rose. 2010. Ergative case and the transitive subject: a view from Nez Perce. *Natural Language & Linguistic Theory* 28(1). 73–120. (doi:10.1007/s11049-009-9081-5)
- DeLancey, Scott. 1981. An interpretation of split ergativity and related patterns. *Language* 57. 626–657.
- Dixon, R.M.W. 1994. *Ergativity*. Cambridge: Cambridge University Press.
- Dixon, Robert M. W. 1983. Nyawaygi. In Dixon, R. M. W & Blake, Barry J (eds.), *Handbook of Australian languages (Volume 3)*, 430–531. Amsterdam: John Benjamins.
- Dryer, Matthew S. 1997. Are grammatical relations universal? In Bybee, Joan L. & Haiman, John & Thompson, Sandra A. (eds.), *Essays on language function and language type: Dedicated to T. Givón*, 115–143. Amsterdam: Benjamins.
- Farrell, Patrick. 2005. *Grammatical relations*. Oxford: Oxford University Press.
- Forker, Diana. 2012. The bi-absolutive construction in Nakh-Daghestanian. *Folia Linguistica. De Gruyter Mouton* 46(1). 75–108. (doi:10.1515/flin.2012.3)
- Haspelmath, Martin. 1990. The grammaticization of passive morphology. *Studies in Language* 14(1). 25–72.
- Haspelmath, Martin. 2011. On S, A, P, T, and R as comparative concepts for alignment typology. *Linguistic Typology* 15(3). 535–567.
- Haspelmath, Martin. 2020. The structural uniqueness of languages and the value of comparison for description. *Asian Languages and Linguistics* 1(2). 346–366. (doi:10.1075/alal.20032.has)
- Haspelmath, Martin. 2021. Role-reference associations and the explanation of argument coding splits. *Linguistics* 59(1). 123–174. <https://doi.org/10.1515/ling-2020-0252>.
- Haspelmath, Martin. 2021. Role-reference associations and the explanation of argument coding splits. *Linguistics* 59(1). 123–174. (doi:10.1515/ling-2020-0252)
- Haspelmath, Martin. 2021a. Role-reference associations and the explanation of argument coding splits. *Linguistics* 59(1). 123–174. <https://doi.org/10.1515/ling-2020-0252>.
- Haspelmath, Martin. 2021b. Towards standardization of morphosyntactic terminology for general linguistics. In Luca Alfieri, Giorgio Francesco Arcodia & Paolo Ramat (eds.), *Linguistic categories, language description and linguistic typology*, 35–57. Amsterdam: Benjamins. DOI: 10.1075/tsl.132.02has, <https://ling.auf.net/lingbuzz/005489>.
- Haspelmath, Martin. 2022. Valency and voice constructions. Leipzig, MPI-EVA, ms. <https://ling.auf.net/lingbuzz/005941>.
- Haspelmath, Martin. 2023. Defining the word. *WORD* 69(3). 283–297. <https://doi.org/10.1080/00437956.2023.2237272>.
- Haspelmath, Martin. 2023. On what a construction is. *Constructions* 15(1). (doi:10.24338/cons-539)
- Haspelmath, Martin. 2024a. Types of clitics in the world's languages. *Linguistic Typology at the Crossroads (to appear)*. <https://ling.auf.net/lingbuzz/007071>.

- Haspelmath, Martin. 2024b. Inflection and derivation as traditional comparative concepts. *Linguistics* (to appear).
- Haspelmath, Martin. 2025. Inverse indexing systems: Beyond transitivity in core argument coding (in prep.).
- Haspelmath, Martin. 2026a. Passive constructions: Universals and language description. (<https://ling.auf.net/lingbuzz/008340>)
- Haspelmath, Martin. 2026b. Valency and voice constructions. Leipzig, MPI-EVA, ms. (<https://ling.auf.net/lingbuzz/005941>)
- Haspelmath, Martin. 2026c. Construction-functions and construction-strategies. In Däbritz, Chris Lasse & Budzisch, Josefine & Basile, Rodolfo (eds.), *Existential and locative constructions*. Berlin: Language Science Press (to appear). (<https://ling.auf.net/lingbuzz/007897>)
- Haude, Katharina & Fernando Zúñiga. 2016. Inverse and symmetrical voice: On languages with two transitive constructions. *Linguistics* 54(3). 443–481. <https://doi.org/10.1515/ling-2016-0009>.
- Heaton, Raina. 2017. *A typology of antipassives, with special reference to Mayan*. United States -- Hawaii: University of Hawai'i at Manoa. (Ph.D.) (<https://www.proquest.com/docview/1953255756/abstract/1269F67B3E1D4F73PQ/1>)
- Heaton, Raina. 2020. Antipassives in crosslinguistic perspective. *Annual Review of Linguistics* 6(1). 131–153. (doi:10.1146/annurev-linguistics-011619-030412)
- Hjelmslev, Louis. 1961. *Prolegomena to a theory of language*. Baltimore: Indiana University Publications in Anthropology and Linguistics.
- Jacques, Guillaume & Anton Antonov. 2014. Direct/inverse systems. *Language and Linguistics Compass* 8(7). 301–318. <https://doi.org/10.1111/lnc3.12079>.
- Janic, Katarzyna & Hemmings, Charlotte. 2021. Alignment shift as functional markedness reversal. *Journal of Historical Linguistics*. John Benjamins 11(2). 299–341. (doi:10.1075/jhl.20017.jan)
- Janic, Katarzyna & Witzlack-Makarevich, Alena. 2021. The multifaceted nature of the antipassive construction. In Janic, Katarzyna & Witzlack-Makarevich, Alena (eds.), *Antipassive: Typology, diachrony, and related constructions*, 1–39. Amsterdam: Benjamins.
- Janic, Katarzyna. 2016. *L'antipassif dans les langues accusatives*. Bruxelles: Lang.
- Johns, Alana. 1992. Deriving ergativity. *Linguistic Inquiry*. The MIT Press 23(1). 57–87.
- Kozinsky, Isaac & Nedjalkov, Vladimir P. & Polinskaja, Maria S. 1988. Antipassive in Chukchee: Oblique object, object incorporation, zero object. In Shibatani, Masayoshi (ed.), *Passive and voice*, 651–706. Amsterdam: Benjamins.
- Kroeger, Paul. 1993. *Phrase structure and grammatical relations in Tagalog*. Stanford, CA: CSLI Publications.
- Kulikov, Leonid I. 2011. Voice typology. In Jae Jung Song (ed.), *The Oxford handbook of linguistic typology*, 368–398. Oxford: Oxford University Press.
- Kulikov, Leonid I. 2011. Voice typology. In Song, Jae Jung (ed.), *The Oxford handbook of linguistic typology*, 368–398. Oxford: Oxford University Press.
- Lazard, Gilbert. 2002. Transitivity revisited as an example of a more strict approach in typological research. *Folia Linguistica* 36(3–4). 141–190. (doi:10.1515/flin.2002.36.3-4.141)
- Lehmann, Christian. 2015. Situation types, valency frames and operations. In Malchukov, Andrej L. & Comrie, Bernard (eds.), *Valency classes in the world's languages: A comparative handbook*, vol. 2, 1547–1595. Berlin: De Gruyter Mouton.
- Mel'čuk, Igor A. & Xolodovič, Aleksandr A. 1970. K teorii grammatičeskogo zaloga. *Narody Azii i Afriki* 1970(4). 111–124.
- Murasugi, Kumiko. 2020. Antipassives. In Siddiqi, Daniel & Barrie, Michael & Gillon, Carrie & Haugen, Jason D. & Mathieu, Eric (eds.), *The Routledge handbook of North American languages*, 275–289. New York: Routledge.
- Nespor, Marina & Irene Vogel. 2007. *Prosodic phonology (2nd edition)*. Berlin: Mouton de Gruyter.

- Newell, Heather, Máire Noonan, Glyne Piggott & Lisa deMena Travis (eds.). 2017. *The structure of words at the interfaces*. Oxford: Oxford University Press.
- Payne, Thomas E. 1982. Role and reference related subject properties and ergativity in Yup'ik Eskimo and Tagalog. *Studies in Language*. John Benjamins 6(1). 75–106. (doi:10.1075/sl.6.1.05pay)
- Polinsky, Maria. 2005. Antipassive constructions. In Haspelmath, Martin & Dryer, Matthew S & Gil, David & Comrie, Bernard (eds.), *The world atlas of language structures*, 438–441. Oxford: Oxford University Press. (<http://wals.info/chapter/108>)
- Polinsky, Maria. 2017. Antipassive. In Coon, Jessica & Massam, Diane & Travis, Lisa deMena (eds.), *The Oxford handbook of ergativity*, 308–331. Oxford: Oxford University Press.
- Polinsky, Maria. 2026. What's so anti-passive about the antipassive? *Total linguistic fact: To honor Michael Silverstein*. (https://www.researchgate.net/profile/Maria-Polinsky/publication/389654060_What's_so_anti-passive_about_the_antipassive/links/67cb7e42e62c604a0dd62b3b/Whats-so-anti-passive-about-the-antipassive.pdf) (Accessed August 9, 2025.)
- Rosen, Carol. 1990. Rethinking Southern Tiwa: The geometry of a triple-agreement language. *Language* 66(4). 669–713.
- Rude, Noel. 1986. Topicality, transitivity, and the direct object in Nez Perce. *International Journal of American Linguistics* 52(2). 124–153. (doi:10.1086/466009)
- Sansò, Andrea. 2017. Where do antipassive constructions come from? *Diachronica* 34(2). 175–218. (doi:10.1075/dia.34.2.02san)
- Sansò, Andrea. 2018. Explaining the diversity of antipassives: Formal grammar vs. (diachronic) typology. *Language and Linguistics Compass* 12(6). e12277. (doi:10.1111/lnc3.12277)
- Say, Sergey. 2005. Antipassive sja-verbs in Russian: Between inflection and derivation. In Dressler, Wolfgang U. & Kastovsky, Dieter & Pfeiffer, Oskar E. & Rainer, Franz (eds.), *Morphology and its demarcations: Selected papers from the 11th Morphology meeting, Vienna, February 2004* (Current Issues in Linguistic Theory 264), 253–275. Amsterdam: Benjamins.
- Schiering, René, Balthasar Bickel & Kristine A. Hildebrandt. 2010. The prosodic word is not universal, but emergent. *Journal of Linguistics* 46(3). 657–709. <https://doi.org/10.1017/S0022226710000216>.
- Seržant, Ilja A. & Janic, Katarzyna Maria & Dermaku, Darja & Ben Dror, Oneg. 2021. Typology of coding patterns and frequency effects of antipassives. *Studies in Language* 45(4). 968–1023. (doi:10.1075/sl.20049.ser)
- Shibatani, Masayoshi. 1985. Passives and related constructions: A prototype analysis. *Language* 61(4). 821–848.
- Shibatani, Masayoshi. 2003. Directional verbs in Japanese. In Erin Shay & Uwe Seibert (eds.), *Motion, direction and location in languages: In honor of Zygmunt Frajzyngier* (Typological Studies in Language). Amsterdam: John Benjamins.
- Song, Jae Jung. 2018. *Linguistic typology*. Oxford: Oxford University Press.
- Spreng, Bettina. 2010. On the conditions for antipassives. *Language and Linguistics Compass* 4(7). 556–575. (doi:10.1111/j.1749-818X.2010.00204.x)
- Tallman, Adam J. R., Sandra Auderset & Hiroto Uchihara (eds.). 2024. *Constituency and convergence in the Americas (to appear)*. Language Science Press. Berlin: Language Science Press. <https://langsci-press.org/catalog/book/291>. (30 November, 2023).
- van Gijn, Rik & Fernando Zúñiga. 2014. Word and the Americanist perspective. *Morphology* 24(3). 135–160. <https://doi.org/10.1007/s11525-014-9242-z>.
- Velázquez-Castillo, Maura. 2008. Voice and transitivity in Guaraní. In Mark Donohue & Søren Wichmann (eds.), *The typology of semantic alignment*, 380–395. Oxford: Oxford University Press. <http://site.ebrary.com/id/10254451>. (14 October, 2013).
- Vigus, Meagan. 2018. Antipassive constructions: Correlations of form and function across languages. *Linguistic Typology*. De Gruyter Mouton 22(3). 339–384. (doi:10.1515/lingty-2018-0013)
- Zúñiga, Fernando & Kittilä, Seppo. 2019. *Grammatical voice*. Cambridge: Cambridge University Press.

- Zúñiga, Fernando & Seppo Kittilä. 2019. *Grammatical voice*. Cambridge: Cambridge University Press.
- Zúñiga, Fernando. 2006. *Deixis and alignment: Inverse systems in indigenous languages of the Americas* (Typological Studies in Language 70). Amsterdam: Benjamins.
- Zúñiga, Fernando. 2007. From the typology of inversion to the typology of alignment. In Matti Miestamo & Bernhard Wälchli (eds.), *New Challenges in typology: Broadening the horizons and redefining the foundations*. Berlin: Mouton de Gruyter. <https://doi.org/10.1515/9783110198904>.
- Zúñiga, Fernando. 2014a. Inversion, obviation, and animacy in native languages of the Americas: Elements for a cross-linguistic survey. *Anthropological Linguistics* 56(3). 334–355.
- Zúñiga, Fernando. 2014b. (Anti-)cliticization in Mapudungun. *Morphology* 24(3). 161–175. <https://doi.org/10.1007/s11525-014-9244-x>.
- Zúñiga, Fernando. 2019. Polysynthesis: A review. *Language and Linguistics Compass* 13(4). e12326. <https://doi.org/10.1111/lnc3.12326>.
- Zúñiga, Fernando. 2023. What is a word? In Carmen Dagostino, Marianne Mithun & Keren Rice (eds.), *The languages and linguistics of Indigenous North America: A comprehensive guide (Vol 1)*, 183–204. Berlin: De Gruyter.