

SSILA conference New Orleans, 2020 January 5

## The polycategoriality parameter: Noun-verb similarities in some languages of the Americas

MARTIN HASPELMATH

*MPI for the Science of Human History (Jena) & Leipzig University*

### I. Big-picture comparison of languages

a widespread **19th century** view (*naive evolutionism*):  
languages of “less advanced” peoples are somehow “primitive”,  
perhaps reflecting different degrees of intellectual development

a widespread **early 20th century** view (*cultural particularism*):  
languages on different continents may be very different from each other,  
partly reflecting the very different cultures of their speakers

a widespread **late 20th century** view (*cognitive universalism*):  
languages from around the world share the same (possibly innate) blueprint  
and vary just slightly from each other

Are we condemned to presenting views of languages that reflect the prevalent ideology in our societies? Or is an **objective perspective** on language structures possible?

#### Reviewer 2:

“Although this abstract states that it does not have any social outcomes or implications, it in fact delves into these indigenous languages in a non-Eurocentric way which is beneficial to the study of indigenous languages. In addition, the nature of this work does benefit the linguistic community. This analysis asks questions that can lead us to more impactful ways to study these languages.”

Using the example of grammatical noun-verb similarities in a range of North American languages, I will argue

- that Reviewer 2 is right that a **non-Eurocentric** perspective is beneficial to the study of non-European languages
- that cognitive universalism may lead us **unwittingly** toward Eurocentrism
- that the noun-verb similarities of some North American languages are indeed striking and worth highlighting as **unusual among the world’s languages**, BUT that they are not exceptions to **robust cultural universals** of grammatical encoding

So Jelinek & Demers (1994) were **not wrong** when they claimed that Salishan languages have “no noun-verb distinction” in some sense (contra Davis et al. 2014: e199).

## 2. Polycategoriality in Salishan, Wakashan, Siouan, Nahuatl, and Mayan

Some North American languages are famous for allowing **action roots** and **thing roots** to occur in predicating function and in referring function **without special function indicators** – i.e. without copulas or relativizers.

(1) Lillooet (= Státimcets, Sł'áł'imxəc; Salishan; Davis et al. 2014: e196)

- a. **action root** → predicating; **thing root** → referring

*ł'iq*      *ta=k<sup>w</sup>úk<sup>w</sup>pi?*=*a*  
arrive      DET=chief=EXIS  
'[The Ø chief] [Ø arrived].'

- b. **thing-root** → predicating

*šmúłáč*      *ta=k<sup>w</sup>úk<sup>w</sup>pi?*=*a*  
woman      DET=chief=EXIS  
'[The Ø chief] [is a woman].'

- c. **action-root** → referring

*k<sup>w</sup>úk<sup>w</sup>pi?*      *ta=ł'iq*=*a*  
chief      DET=arrive=EXIS  
'[The one who arrived] [is a chief].'

(2) Nuuchahnulth (Wakashan; Swadesh 1939: 78)

- a. *Mamo:k-ma*      *qo:ʔas-ʔi*.

work-PRS.3SG man-DET  
'[The Ø man] [is working].'

- b. *Qo:ʔas-ma*      *mamo:k-ʔi*.

work-PRS.3SG work-DET  
'[The one who is working] [is a man].'

(3) Lakhota (Siouan; Van Valin 2019: 4)

- a. *Winyan* *kiŋ* *hená*      *wačhí-pi-kte-šni*.

woman DEF those dance-PL-FUT-NEG  
'[Those Ø women] [will Ø not dance].'

- b. *Wačhí-pi* *kiŋ* *hená*      *winyan-pi-kte-šni*.

dance-PL DEF those woman-PL-FUT-NEG  
'[Those dancers] [will not be women].'

(4) Classical Nahuatl (Uto-Aztecán; Launey 1994)

- a. *Tzàtzi* *in* *pilli*.

cry      DEF child  
'[The Ø child] [Ø cries].'

- b. *Pilli*      *in*      *tzàtzi*.

child      DEF cry  
'[The one who cries] [is a child].'

(5) Yucatec Maya (Vapnarsky 2013: 44)

- a. *T-u-hantah le kay-o'.*  
 CPL-3.ERG-eat DEF fish-DET  
 'She [Ø eats] [the Ø fish].'
- b. *Kay le t-u-hantah-o'.*  
 fish DEF CPL-3.ERG-eat-DET  
 '[Fish is] [what she eats].'

How are these languages different from English and other European languages?

(5) The polycategoriality parameter

Value A: **Thing roots** require a **copula** in predicating function, and **action roots** require a **relativizer** in referring function.

Value B: **Thing roots do not** require a copula in predicating function, and **action roots do not** require a relativizer in referring function.

Value B is a kind of “**marker economy**” – these languages **lack function indicators** where other languages have them.

The **propositional act function** (predicating vs. referring, cf. Croft 2000) remains implicit and is inferred from the context (e.g. the word order).

Marker economy is frequently found elsewhere, because many grammatical meanings can be inferred from the context, e.g.

### **number marker economy**

many languages do not make a singular-plural distinction in nouns, especially when combined with numerals

cf. Haspelmath (2005) in *WALS*

### **tense marker economy**

many languages do not make a present-future distinction in verbs, especially when combined with time adverbials

cf. Dahl & Velupillai (2005) in *WALS*

### **evidential marker economy**

many languages do not make an evidential distinction in verbs

cf. de Haan (2005) in *WALS*

Thus:

Lillooet, Nuuchahnulth, Lakhota, Nahuatl and Yucatec exhibit **function indicator economy**, of a kind that one might call “polycategoriality” (cf. Vapnarsky & Veneziano (eds.) 2017).

But does this mean that there is “no noun-verb distinction” in these languages?

### 3. Language-particular classes

What does it mean to say that a language “has a noun-verb distinction”? Clearly, it makes a **general statement** about the language, which goes beyond its own language system – it is a comparative claim.

But morphosyntactic word classes are **language-particular**, created in order to describe a language in a complete way with respect to its morphosyntactic constructions.

e.g. Maricopa (Gordon 1986: 29; 90):

**Pluralizables:**

*humar/humaar* ‘child(ren)’  
*mhay/mhaa* ‘boy(s)’  
*‘iipaal’iipash* ‘man/men’  
*hmii/hshmee* ‘be tall’  
*uukup/shuukuup* ‘dig’  
*nak/anaak* ‘sit’

**Nonpluralizables:**

*avhay* ‘dress’  
*shviily* ‘feather’  
*vii* ‘rock’  
*hly’a* ‘month’  
*kwnho* ‘basket’  
*nyik* ‘rope’

e.g. Godoberi (Kibrik (ed.) 1996)

**Gendered Words:**

*(b-)ičǎ-* ‘sell’  
*q’aruma(-b)* ‘greedy’

**Genderless Words:**

*bit’i-* ‘tear’  
*mik’isi-* ‘young’

e.g. English

**Much-Premodified Words:**

*money*  
*admire*  
*regret*

*We don’t have much money.*  
*We much admire your technique.*  
*We much regret the inconvenience.*

**Non-Much Words:**

*pay*  
*love*  
*hate*

*\*We don’t much pay.*  
*\*We much hate your uncle.*  
*\*We much love our teacher.*

e.g. French

**Article Person Words**

*la mère* ‘the mother’  
*le frère* ‘the brother’

**Articleless Person Words**

*Hélène*  
*Jacques*

In what way can such **language-particular classes** be associated with **general classes**? This is not clear – the literature usually focuses on the labeling question, e.g. Schachter (1985: 7):

“The label **noun** is assigned to the class of words in which occur the names of most persons, places and things.”

This is indeed how we often **label** our classes – and if there is no class called “noun”, we often create one by lumping some classes.

(e.g. “Common Noun”, “Proper Noun”)

But sometimes we do not follow this advice for labeling –  
cf. Maricopa, where we DO NOT call the Nonpluralizables “nouns”,  
even though most thing roots are in that class:

**Pluralizables:**

*'iipaal/iipash* ‘man/men’  
*hmii/hshmee* ‘be tall’  
*uukup/shuukuup* ‘dig’  
*nak/anaak* ‘sit’

**Nonpluralizables:**

*vii* ‘rock’  
*hly'a* ‘month’  
*kwnho* ‘basket’  
*nyik* ‘rope’

Regardless of the labeling, how can we tell in general whether there is a “noun-verb distinction” in a language?

#### 4. Why the “noun-verb distinctness question” cannot be answered

Many linguists have formulated the issue concerning noun-verb similarities as follows:

Does language X have a noun-verb distinction?

cf. Bloomfield (1933)  
Lyons (1977)  
Chung (2012)  
Davis et al. (2014)  
... and many others

But in practice, this question cannot be answered in a way that does not already presuppose the answer.

To answer it in an empirical way, we would need either (A) or (B):

(A) a way to identify “noun” and “verb” **by the same positive criteria way in all languages**

(B) complete knowledge of **all the possibly relevant morphosyntactic constructions** of the language

Ad (A):

- Since each language has **different morphosyntactic constructions**, “noun” and “verb” cannot be identified uniformly in all languages – unless we **limit the choice of valid criteria** to those that are potentially relevant in all languages.

- Features such as articles, gender, number, tense, aspect are very variable across languages and cannot be used to identify nouns and verbs across languages.

- (• Only **function indicators** might be a possible basis for comparison: **Copulas** are characteristic of nouns in predicating function, and **relativizers** are characteristic of verbs in referring function.)

Ad (B):

- One might object: No, other criteria are valid to establish a noun-verb distinction as well – we just need to look for them **very hard**, because the distinction is often “**subtle**”.
- BUT: We can never be sure that we have examined **all** the potentially relevant constructions across languages – which means in practice that it is not possible to demonstrate the lack of a noun-verb distinction.

“absence of evidence is not necessarily evidence of absence (of a noun-verb distinction)” (Davis et al. 2014: e197)? – but in practice, on this view, the absence of a noun-verb distinction cannot be demonstrated

- MOREOVER: One could always claim that an alleged distinguishing criterion **merely establishes subclasses**.

So can we take the fact that the Verbs of Straits Salishan can follow Auxiliaries as “evidence for a noun-verb distinction” (cf. Montler 2003)?

NO – because one could always say that Straits Salishan has several subclasses of its “Full Words” class.

In fact, one might even claim that English has a class of “Full Words”, characterized by the possibility of a **Number suffix -s** (with Plural meaning in one subclass, e.g. *cat-s*, and Singular meaning in another subclass, e.g. *eat-s*)

Thus, all we can do, in practice, is ask the following questions:

(A) Are there **universal tendencies** in the way languages treat thing roots and action roots in referring function and in predicating function?

(B) Are there **salient differences** between languages in this respect?

We already saw that there is a salient difference –

the **polycategoriality parameter**.

It is not clear to me that this reflects any kind of really “deep” difference between languages (cf. Launey’s “omnipredicativity”), but it is certainly “salient” – and apparently cross-linguistically uncommon.

Absence of a copula is common – cf. Stassen (2005) in WALS.

But absence of a relativizer does not seem to be common (though I do not know of any systematic cross-linguistic research).

## 5. Cultural universalism and cognitive universalism

(A) Are there **universal tendencies** in the way languages treat thing roots and action roots in referring function and in predicating function?

YES (Croft 1991; 2000):

Languages show a very strong tendency to have **function indicators** for those functions that are not usually associated with semantic root types:

- relativizers for action roots in referring function
- copulas for thing roots in predicating function

[WHO arrives first] gets a prize.	(relativizer)
Mother [WAS a linguist].	(copula)

Languages do not have relativizers for the thing roots in referring function, or copulas for action roots in predicating function:

*[WHO mother] left for a conference.	(relativizer)
*Mother [WAS arrived first].	(copula)

This generalization is part of a larger universal:

### (2) The grammatical form-frequency correspondence universal

When two grammatical construction types that differ minimally occur with significantly different frequencies, the **less frequent** construction tends to be **overtly coded**, while the **more frequent construction** tends to be **zero-coded** (Haspelmath 2020).

Function indicators occur in the less frequent construction types, i.e. when thing roots occur predicatively or action roots occur in referring function.

This universal is a **cultural universal**, because it has a good **functional-adaptive explanation**: Languages use short forms for frequently occurring meanings because speakers try to expend only as much energy as they must (Zipf 1935; Gibson et al. 2019).

Cultures vary widely across human populations, and most anthropologists study cultural peculiarities – but whenever aspects of cultures are universal, they seem to have functional-adaptive explanations.

By contrast:

**Cognitive universalism** starts out with the idea that similarities between languages are due to a biocognitive blueprint – an innate toolkit of categories (“UG”).

If there is a small set of innate categories, then it makes sense to use the Uniformity Principle as a heuristic:

(7) Uniformity Principle

“In the absence of compelling evidence to the contrary, assume languages to be uniform, with variety restricted to easily detectable properties of utterances.”

(Chomsky 2001: 2)

(But there is no good reason to assume that there is a rich innate toolkit of categories – most language universals can be explained without such a strong assumption, which is implausible anyway for biological reasons, cf. Berwick & Chomsky 2016).

If one adopts the strong biocognitive toolkit assumption as a leading assumption, it is easy to “find evidence” for it – any morphosyntactic property that remotely looks like a “noun-verb” distinction can be taken as evidence for it.

The “noun-verb” distinction can look very different in different languages (e.g. Nonpluralizable/Pluralizable in Maricopa), but one could still say that **all languages “have a noun-verb distinction”** (even if it is very “subtle”).

But this carries a danger of **Eurocentrism**, because in practice, it implies that we should look for those categories that happen to be salient in our current theories, which tend to be strongly influenced by English (plus Latin, German, Spanish and other influential languages).

Each claim that “the distinction between X and Y is universal” is **virtually impossible to prove wrong**, because one cannot show (in practice) that a language completely lacks the distinction.

(See Haspelmath 2012 on Chung 2012, which presents a rather Eurocentric view of Chamorro word classes.)

A Maricopa linguist could claim:

English is **an example of category neutrality**, because it lacks the important distinction between Pluralizables and Nonpluralizables – the counterparts of both classes are pluralizable in English:

e.g. Maricopa (Gordon 1986: 29;90):

**Pluralizables:**

*humar/humaar* ‘child(ren)’

*mhay/mhaa* ‘boy(s)’

*‘iipaal/‘iipash* ‘man/men’

*hmii/hshmee* ‘be tall’

*uukup/shuukuup* ‘dig’

*nak/anaak* ‘sit’

**Nonpluralizables:**

*avhay* ‘dress’

*shviily* ‘feather’

*vii* ‘rock’

*hly’a* ‘month’

*kwnho* ‘basket’

*nyik* ‘rope’

This would be a Maricopa-centric view of the world, though a Maricopa cognitive universalist could point out that if you look harder in English, you do find some words that are not pluralizable (e.g. *milk*, *rice*, *should*) – so this would be evidence for the cognitive universality of the two Major Parts of Speech after all.

## 6. Conclusion

Jelinek & Demers (1994) were not completely wrong, and there was no reason for Eloise Jelinek to “recant” – Salishan languages are indeed strikingly different from English and other European languages.

If one allows for cultural universals in addition to bicognitive universals, then there is no reason to start out with the **assumption** of cognitive universals.

If universality is driven by functional-adaptive forces, then the peculiarities of different cultures – the ways in which they are saliently different from each other – can come into focus again.

This allows us to combine **cultural universalism** (which asserts the unity of humankind) with **cultural particularism** (which asserts the value of the peculiarities of specific cultures).

We need not deny *cognitive universalism* either, but it does not seem to play a big role in grammar. However, this is an empirical question – if cognitive universalism is treated **not as an assumption, but as a testable claim**, we might well find evidence for it eventually.

In practical terms, this means that **the search for “subtle”, “hidden” evidence for a general category is meaningless**. Languages make many subtle distinctions, and we should document them, but they cannot be evidence for general categories of Human Language.

## References

- Berwick, Robert C. & Noam Chomsky. 2016. *Why only us: Language and evolution*. Cambridge, MA: MIT Press.
- Bloomfield, Leonard. 1933. *Language*. New York: H. Holt and Company.
- Chomsky, Noam A. 2001. Derivation by phase. In Michael Kenstowicz (ed.), *Ken Hale: A life in language*, 1–52. Cambridge, MA: MIT Press.
- Chung, Sandra. 2012. Are lexical categories universal? The view from Chamorro. *Theoretical Linguistics* 38(1–2). 1–56. doi:[10.1515/tl-2012-0001](https://doi.org/10.1515/tl-2012-0001).
- Croft, William. 2000. Parts of speech as language universals and as language-particular categories. In Petra M. Vogel & Bernard Comrie (eds.), *Approaches to the typology of word classes*, 65–102. Berlin: Mouton de Gruyter.
- Croft, William. 2001. *Radical construction grammar: Syntactic theory in typological perspective*. Oxford: Oxford University Press.
- Dahl, Östen & Viveka Velupillai. 2005. The future tense. In Martin Haspelmath, Matthew S Dryer, David Gil & Bernard Comrie (eds.), *The world atlas of language structures*, 270–. Oxford: Oxford University Press. <https://wals.info/chapter/67>.
- Davis, Henry, Carrie Gillon & Lisa Matthewson. 2014. How to investigate linguistic diversity: Lessons from the Pacific Northwest. *Language* 90(4). e180–e226.

- Gibson, Edward, Richard Futrell, Steven T. Piandadosi, Isabelle Dautriche, Kyle Mahowald, Leon Bergen & Roger Levy. 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).
- Gordon, Lynn. 1986. *Maricopa morphology and syntax*. Berkeley: University of California Press.
- Haan, Ferdinand de. 2005. Coding of evidentiality. In Martin Haspelmath, Matthew S. Dryer, David Gil & Bernard Comrie (eds.), *The world atlas of language structures*, xxx. Oxford: Oxford University Press. <https://wals.info/chapter/78>.
- Haspelmath, Martin. 2005. Occurrence of nominal plurality. In Martin Haspelmath, Matthew S. Dryer, David Gil & Bernard Comrie (eds.), *The world atlas of language structures*, xxx. Oxford: Oxford University Press. <https://wals.info/chapter/34>.
- Haspelmath, Martin. 2012. Escaping ethnocentrism in the study of word-class universals. *Theoretical Linguistics* 38(1–2). doi:[10.1515/tl-2012-0004](https://doi.org/10.1515/tl-2012-0004).
- 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. <https://zenodo.org/record/3519206>.
- Haspelmath, Martin. 2020. Explaining grammatical coding asymmetries: Form-frequency correspondences and predictability. *to appear*. <https://ling.auf.net/lingbuzz/004531>.
- Jelinek, Eloise & Richard A. Demers. 1994. Predicates and pronominal arguments in Straits Salish. *Language* 70(4). 697–736. doi:[10.2307/416325](https://doi.org/10.2307/416325).
- Kibrik, Alexandr (ed.). 1996. *Godoberi*. München: Lincom Europa.
- Launey, Michel. 1994. *Une grammaire omniprédicative: Essai sur la morphosyntaxe du nahuatl classique*. CNRS Editions.
- Lyons, John. 1977. *Semantics*. Cambridge: Cambridge University Press.
- Montler, Timothy. 2003. Auxiliaries and other categories in Straits Salishan. *International Journal of American Linguistics* 69(2). 103–134. doi:[10.1086/379680](https://doi.org/10.1086/379680).
- Schachter, Paul. 1985. Parts-of-speech systems. In Timothy Shopen (ed.), *Language typology and syntactic description, vol. 1.*, 3–61. Cambridge: Cambridge University Press.
- Stassen, Leon. 2005. Zero copula for predicate nominals. In Martin Haspelmath, Matthew S. Dryer, David Gil & Bernard Comrie (eds.), *The world atlas of language structure*. Oxford: Oxford University Press. (<http://wals.info/chapter/120>).
- Swadesh, Morris. 1938. Nootka internal syntax. *International Journal of American Linguistics* 9(2/4). 77–102. doi:[10.1086/463820](https://doi.org/10.1086/463820).
- Van Valin, Robert D. 2019. On the status of lexical categories in RRG. <https://ubwp.buffalo.edu/rrg2019/wp-content/uploads/sites/101/2019/09/Van-Valin.pdf>.
- Vapnarsky, Valentina. 2013. Is Yucatec Maya an omnipredicative language? Predication, the copula and focus constructions. *STUF - Language Typology and Universals Sprachtypologie und Universalienforschung* 66(1). 40–86. doi:[10.1524/stuf.2013.0004](https://doi.org/10.1524/stuf.2013.0004).
- Vapnarsky, Valentina & Edy Veneziano (eds.). 2017. *Lexical polycategoriality: Cross-linguistic, cross-theoretical and language acquisition approaches*. Amsterdam: Benjamins.
- Zipf, George Kingsley. 1935. *The psycho-biology of language: An introduction to dynamic philology*. Cambridge: MIT Press.