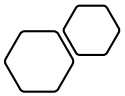
Linguistic Typology for NLP researchers: Methods and Resources in the 21st century



Linguistic analysis: morphemes, words, clauses, IGT, grammars and comparative concepts

Luigi Talamo (Saarland University)

Credits

The present session closely follows the first two chapters of 'William Croft, *Morphosyntax: constructions of the world's* (Croft 2022) *languages.* Cambridge: Cambridge University Press. (PDFs courtesy of prof. Croft)

In these two sessions (or one session and half) we will deal with linguistic analysis, focusing on the morphosyntactic level. Croft 2022 is a recent handbook on morphosyntax from a typological perspective.

Chapter 1: *Grammatical Constructions, Semantic Classes and Information Packaging*

Introduction: What is Morphosyntax

- It is the combination of **Morphology and Syntax**, two different levels of analysis.
- Morphology: internal structure of the word. It has to do with things like morphemes i.e., prefixes, suffixes, infixes, stem, root and studies the formal and functional shape (morpho-) of words;
- Syntax: internal structure of the sentence/utterance. It has to do with things like words, more specifically, how words formally and functionally combine in a sentence.

More than often, grammatical constructions involve both morphology and syntax, hence the need for morpho-syntax.

Morphology: upper limit is the word, lower limit morphemes
Syntax: upper limit is the sentence/utterance, lower limit word
Being a combination of both levels of analysis, we can infer that Morphosyntax studies the internal structure of the sentence, ``both above the word level and below it'' (Croft 2022:2)
You will often hear in this class of constructions and how they are central for our discuss; I'll explain this later in more detail, for the moment keep in mind that constructions are regarded as the proper units for the grammatical analysis.

Introduction: What is Morphosyntax

English numeral constructions

One tree

Two tree-s

Three tree-s

This is the first of the several constructions that we will encounter in our class and compare with similar constructions around the world. It exemplifies how morphology and syntax often work together in constructions: can you tell me how?

Introduction: What is Morphosyntax

English numeral modification

One tree

Two tree-**s**

Three tree-**s**

Morphology: 0/-s morpheme indicating the plural (inflection)

Syntax: juxtaposition of the numeral with the quantified object, order of number and noun (word order)

Introduction: What is Morphosyntax

There are some constructions involving only…

- Morphology:

inflection (I play/he play-s)

derivation (play -> play-ful)

- Syntax:

word order (black cat and not *cat black)

…but we won't discuss them in this class, which is devoted to Morphosyntax.
Morphology here will be seen as ancillary to Syntax, serving it in constructions.

Introduction: What is Morphosyntax

Grammaticalization processes: ”Today’s morphology is yesterday’s syntax” (Tom Givòn)

Morphology and **syntax** are strongly tied from a **diachronic perspective**

Yesterday: I **am** happy to be a teacher.

Today: I’**m** happy to be a teacher.

Another reason to study morphology and syntax together is that they are often connected in diachronic cycle i.e., bound morphemes such as ‘m originates in the word (syntax) am.
This will be useful in our world travel, as some languages use multiword strategy for functions or meanings that in other languages are covered by a single word.

Introduction: Anatomy of a Morphosyntactic construction

The basic structure of a construction is articulated into

- Form: **Morphosyntax**
- Function: **Information packaging** and **meaning**

So, it's not all about **meaning**! Morphosyntactic constructions also differ in **which type of information they package**. Compare:

Fifteen trees vs. Trees **number fifteen**

A construction is conventionally given within square brackets:

English Numeral Modification: [NUM NOUN-NUMBER]

Like all Saussurean signs, the structure of a construction is articulated into form and function. We have already seen how form is represented here by Morphosyntax, which in turn amounts to morphology plus syntax. In turn morphemes amount to phonemes and phonemes to phones produced by our phonatory system. But this is another story.

Let's focus now on function, which is further articulated into two components: information packaging and meaning.

Information packaging refers to the how the information is communicated: in the first sentence, I convey the meaning that trees are modified by the numeral fifteen, while in the second sentence `I predicate the fact that trees are fifteen (I could have also said something like Trees are fifteen). The meaning is the same (trees are 15), but is conveyed in different ways, hence the information packaging.

Introduction: theoretical framework

We will work in a **constructionist and functional framework** of Linguistics.

1. The **construction** is the proper unit for grammatical analysis (CG)
2. **Form follows function** i.e., the analysis shall explain how Morphosyntactic form expresses function (CG)
3. The analysis shall be conducted from **a cross-linguistic perspective** i.e., which morphosyntactic constructions express the same function **across languages**?

The first two assumptions are shared by constructionist models of language theory, such as Construction Grammar, Radical Construction Grammar and Role and Reference Grammar, while the third is the hallmark of Linguistic Typology, a linguistics discipline studying the diversity of world languages and whether general principles such as statement, implication and hierarchy govern such diversity. An example of an implication is 'if a language has the gender, then it has the number".

Introduction: word classes vs. semantic classes

Grammars and descriptive works in Linguistics make a lot of reference to **word classes.**

Word classes such as adjective, noun, verb and adverb, have been described from a **Eurocentric perspective** that is, we apply categories defined on English, Latin and German to other languages.

This results in **confusing and non-consistent usages** of these terms, as in the four definitions discussed on page 5 of Croft 2022.

Semantic classes, such as **property**, **object** and **action concepts**, can be instead used to study and describe languages from a **cross-linguistic perspective**, as they are based on **semantics**, which is universal.

Back in the old days, missionaries sometimes wrote grammars of exotic languages using the Latin grammar as a template. This resulted in weird statement like 'Language X does not have the vocative or the genitive case'. The grammar of European languages (which is somewhat averaged to the grammar of English in this course) is just a piece of the Universal Grammar: we will learn that many languages of the world lack a lot of European grammatical stuff, in turn showing things that we do not find in English (and in other European languages).

Don't be bothered for now by definitions given on page 5: we will see these definitions and meet the relative languages later.

We should distinguish here between word classes and semantic classes, using for instance adjectives for word classes (syntactic categories) and property concepts for semantic classes.

Introduction: word classes and constructions

Let's start with the **form** of word classes.

Word classes are **defined as words** occurring in **constructions, filling a particular role.**

For instance, how do we state that 'tall is an English Adjective'? Because this word occurs in the following morpho-syntaxtic constructions: (the form)

1. **As modifiers of noun:** a tall tree
2. **As the complement of a copula *be* in predication**: *That tree is tall*
3. **In a morphological construction** (**inflection**): *tall-er, tall-est*
4. **As modified by degree expressions**: *very tall, a little tall*

Does it look obscure? A bit circular? Don't worry, we simply define word classes as collection of constructions (syntactic patterns). In the example, the English Adjective word class is a collection of four different constructions. Tall is an English Adjective, since it follows the same syntactic pattern of other adjectives such as short, quick, blue, old, …

Why language-specific? Because it may be the case that constructions 1-4 do not apply to other languages. For instance, do they apply to German Adjectives?

Introduction: word classes and constructions

Syntactic categories like word classes are defined by their **occurrence** in constructions:

This is not restricted to **syntax**!

Morphological categories such

- as Third Person Singular Present are also defined by constructions: [_-s], in which –s fills a specific role in a construction
- Derivation of Adjectives from Noun: [_-ful]

Introduction: word classes and constructions

A word class (or another syntactic category) corresponds to just one construction. Or, in other words, **on the formal side** the four constructions define **four different word classes**.

Then, are there four adjectival classes in English? Strictly speaking, yes.

1. An alive insect ❌
2. The insect is alive. ✅
3. Ants are aliver/more alive than bees. ❌
4. The insect is very alive. ❌

Do all English Adjectives fulfil all the four constructions? In other words, is English Adjective a word class defined by the four constructions?
We assumed that English Adjective is a unique class, which is defined by 4 constructions. Actually, this is not. Rather, it's a class defined by the overlapping of four minor classes.
Some adjectives belong to the four classes, others to just two or three. The adjective 'alive' belongs to just one adjectival class. It's a matter of gradience!

Introduction: word classes and constructions

If word classes and syntactic/morphological categories amount to word/unit/morpheme filling a specific **constructions in specific languages**, how do we contrast a language (for instance, English) with other languages?

Well, the **form** of a construction is not enough. We need the **function**:

- Adjectives? Give the **functional equivalent** of the constructions {1,2,3,4}
- Third Person Singular? Give the **functional equivalent** of the constructions [_-s]
- Adjective from Noun? Give the **functional equivalent** of the constructions [_-ful]

You can think of functional equivalent as the best translation you can find: remember, you do not have to translate only the meaning, but also how the information is packaged! English Adjective {1,2,3,4}

Introduction: Why information packaging?

Let's now take a look to the **function** of the construction: **information packaging** + **meaning**.

Three major types of information packing aka **propositional acts:**

1. **reference -** what the speaker is talking about
2. **predication -** what the speaker is asserting about the referents in a particular utterance
3. **modification -** additional information provided about the referent

(Searle 1969, Croft 1991, Croft 2001, Croft 2022:11)

The same **semantic content** (meaning) can be given in **three different ways**.

We have seen that morpho-syntactic constructions can be compared cross-linguistically (i.e., English vs. other languages) by their function. We have also seen that function is further articulated into content/meaning and information packaging, or how the meaning/content is presented from the speaker to the hearer(s).
Propositional acts go back to the work of the language philosopher Searle and Croft 1991's further elaboration.
As with the fifteen trees example, we can use the meaning of a word in three different and basic way. Take for instance the word tall. In 'the tall tree', we use this word to adding additional information on this referent by expressing the property of tallness: we may need this information, say, to distinguish the tree from other trees: the tall tree, not the bushy one.
If we say instead something like 'that tree is tall', we are predicating the fact that the referent has a given property: we can use this information packaging type i.e., predication to argue that a tree is too tall for being climbed, or the like. Is there another way to convey the meaning tall? Yes. It's not so used, but sometimes we need to talk about the fact of being tall i.e., tallness. For instance, the tallness of that tree can be risky for its stability.

Introduction: Why information packaging?

Three major types of **semantic content** (meaning) aka **semantic classes**:

- **Object** i.e., concepts like car, tree, woman, love;
- **Action** i.e., concepts like drive, climb, live, (to) love;
- **Property** i.e., concepts like fancy, tall, beautiful, loving.

Again: are semantic classes enough to define word classes (= a type of construction)?

We can arrange semantic content into three different types: object, action and property. Car, tree, woman are NOUNS in English, so (object = nouns). Drive, climb, live are VERBS in English, so (action = verbs). Fancy, tall, beautiful are ADJECTIVES in English so (properties = adjectives). But what about words like love/ to love/ loving?

Introduction: Why information packaging?

Traditional word classes like **nouns, adjectives and verbs** can be re-defined by using both **propositional acts** and **semantic classes**.

Semantic class	**Propositional act**		
	reference	***modification***	***predication***
object	nominal phrase *head:* noun	possessive modifier/ genitive phrase	predicate nominal
property	property referring phrase	adjectival phrase *head:* adjective	predicate adjectival
action	complement (clause)	relative clause	verbal clause *head:* verb
(all semantic classes)	referring/argument phrase *head:* referent expression	attributive phrase *head:* modifier	clause: *head:* predicate

Croft 2022:30

As we have done before with the TALL concept (the English word tall, the German word groß, the French word haut, the Greek word ψηλό), we can construct cross-linguistically valid word classes by combining semantic class and propositional acts. We have 9 possible combinations, as shown in the picture taken from Croft 2001, a book introducing the framework of Radical Construction Grammar.
Croft observes that cross-linguistically three combinations are more frequent than others: these are given in CAPITALS and correspond to the traditional word classes of nouns, adjectives and verbs.

Introduction: Why information packaging?

We have **three different levels** here:

- **Semantic class**: object, action and property (UNIVERSAL)
- **Propositional act**: reference, modification and predication (UNIVERSAL)
- **Word class**: constructions like noun, adjective and verb (but also numerals, demonstratives, prepositions, …). These are language-specific and defined by different constructions in different languages. (LANGUAGE-SPECIFIC)

Each language of the world differently maps the schema given before to its **specific constructions**: we have English Nouns, Swedish prepositions, Italian verbs, Armenian Adjectives…

Croft's schema shall be universal, however only some languages of the world (apparently) show a neat distinction between these nine constructions. EU languages are among these languages, but the case of English adjectives show that things are actually more complicated.

Let's meet then four extra-European languages: Acehnese (/ˌɑːtʃəˈniːz/: Austronesian), Inupiat (Eskimo-Aleut), Sabane (**Nambiquaran) and Sidaama (Afro-Asiatic).**

Introduction: Why information packaging?

- "Sidaama numerals are adjectives" (Kawachi 2007:135)

Sidaama **numerals** occur in the **same modification construction** used for **prototypical modifiers**.

- "Acehnese has no class of adjectives" (Durie 1985:101)

Acehnese does not have **a modification construction** used specifically for **property words**.

- "Iñupiaq numerals are nouns" (Lanz 2010:110)

Iñupiaq **numerals** occur in the **same reference function** used for **prototypical referents**.

By combining semantic classes and propositional acts, we can rephrase the definitions given on page 5, as we have done with the nine combinations. Traditional definitions as given by grammars and descriptive works often confound the information packaging, the semantic class and the word class level.

Let's start with our Afro-Asiatic language, Sidaama: does the definition mean that numerals (numbers, 1,2,3,4,…) are property concepts like tall and beautiful? Well, not really. They show the same modification construction as adjectives, but they are not property concepts, they are still numerals from a semantic perspective.

The case of Acehnese is even worse: does this language lack concepts like tall, beautiful and loving (semantic class) or miss the modification function i.e., we cannot add information to a referent? Again, no. Rather, it means that the modification construction is not exclusively used for property concepts, as in our languages.

Finally, the Eskimo-Aleutin language is similar to Sidaama: only, nouns and not adjectives behave like numerals. Smth like '15 is the number of trees/eggs/dogs'

Introduction: Why information packaging?

Sabane: “The morphemes **ano** and **wola/wolata** are used to express plurality, meaning ‘much/many/a lot’...
ano is a **quantifier**, while **wola(ta)** is an **adverb**” (Antunes de Araujo 2004:96)

naysunum-ka **ano**-it-al-i
land-OBJ much-VS-PRS.N-ASSR
‘There is plenty of land.’

wolata amayl-i-al-i
a_lot_more to_rain-VS-PRS.N-ASSR
‘It is raining very hard.’

In Sabane the two morphemes ano and wolata are quite synonymic . Ano is defined using a semantic class (quantifiers are concepts like many, a lot, few, etc.), while the second one is either a word class (Sabane adverb) or a propositional act (adverbs are modifier of verbs). So, both ano and wolata occur in a construction of a quantifier used with a modification function.

Introduction: How do we compare constructions within and across languages?

- We use comparative concepts, as defined by Haspelmath (2010):

"Theoretical concepts used for cross-linguistic comparison"

- Concepts belonging to semantic classes and – better – propositional acts can be used as comparative concepts for the functional part of the construction. Even better, we can speak of **property reference construction**, **object reference construction** or **action modification construction**. (WHAT)
- But what about **the form of a construction**? (HOW)

We need to search for the **strategy** that expresses the construction.

We compare English morphosyntactic constructions by using their function and not only their semantic class or, worse, their morphosyntactic form.
We introduce here a term from another linguist, this time a European (German) one, Martin Haspelmath. Much like Croft, he proposes to use semantic and/or function to compare languages and not language-specific things like adjective or nominative case.
In order to contrast English with other languages, we then search for the pairing of function (semantic class + propositional act) and morpho-syntax: we can refer to these comparative concepts by using their functional combinations: property reference construction, et cetera. However, this is only part of the story. Property reference construction or similar expressions do not tell us HOW English and other languages express this function.

But how do we compare the other side of the construction, its form? We need another type of comparative concept, a strategy.

Introduction: How do we compare constructions within and across languages?

- **strategy**: a construction in a language (or any language), used to express a particular combination of semantic content and information packaging (the 'what'), that is further distinguished by certain characteristics of grammatical form that can be defined in a crosslinguistically consistent fashion (the 'how'). (Croft 2022:17)

For instance, we can speak of **inflection strategy, noun modification strategy or degree modification strategy.**

- The same construction **can be expressed by different strategies across and within languages**. (cf. the English adjective)

You can think of strategies as the formal part of the construction: they can be (partially) defined in a cross-linguistic fashion.

Introduction: How do we compare constructions within and across languages?

Three types of strategies:

- **Encoding strategy**: the strategy expresses (=encodes) the same function in different languages. Within a language, we can have different encoding strategies for the same function.
- **System strategy**: the strategy is defined in terms **of the similarities and differences in the forms** of different constructions. (SYNCHRONY)
- **Recruitment strategy**: over time, the strategy **has been extended beyond its original construction**, covering the form of other constructions. (DIACHRONY)

Introduction: How do we compare constructions within and across languages?

Encoding strategy:

- across languages: same function in different languages. For instance verbal copula in English and Italian;
- within a language: different encoding strategies for the same construction. For instance,

English Object Modification (function)
strategy **a.** the regulations of the university
strategy **b.** the university's regulations
strategy **c.** university regulations

Introduction: How do we compare constructions within and across languages?

System strategy: the strategy is defined in terms **of the similarities and differences in the forms** of different constructions. (SYNCHRONY)

Example: the accusative nominative alignment and its nominative strategy (NOM), which appears in two constructions: transitive and intransitive.

Turkish (Turkic)
Adam pastay-ı yiyor.
man.**NOM** cake-**ACC** eats
'The man eats the cake'.

Adam koşuyor.
man.**NOM** runs
'The man runs.'

German (Indo-European)
Der Mann isst den Kuchen.
the.**NOM** man.**NOM** eats the.**ACC** cake.**ACC**
'The man eats the cake.'

Der Mann läuft.
the.**NOM** man.**NOM** runs
'The man runs.'

A system strategy is defined in synchronic term by looking from a cross-linguistic perspective to different strategies encoding similar functions. Take for instance the system of alignment we encounter in European languages, i.e. accusative-nominative alignment.

Let's try to give a comparative definition for NOM and ACC. Remember, you cannot use morpho-syntactic terms like subject and object, as they need a definition and may be not universal. You can use semantic concepts like 'the one who performs the action' (agent) or 'the one who undergoes the action' (patient), and so on. Also, you can use pre-theoretical concepts like 'the only argument'.

In both German and Turkish we can define the nominative strategy as the only argument of an intransitive sentence, which is also used as one of the two arguments of a transitive sentence, in this case performing the action. We have defined the nominative strategy using two different constructions: 'the only argument of an intransitive sentence' and 'one of the two arguments of a transitive sentence, in this case performing the action'.

Introduction: How do we compare constructions within and across languages?

System strategy: the strategy is defined in terms **of the similarities and differences in the forms** of different constructions. (SYNCHRONY)

Example: the ergative-absolutive system and its absolutive (ABS) strategy, which appears in two constructions: transitive and intransitive.

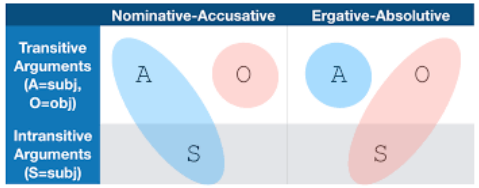
Yuwaalaraay (Pama-Nyungan, Williams 1980:36)
duyu-gu nama dayn-Ø yi:-y
snake-**ERG** that man-**ABS** bite-NFUT
'The snake bit the man.'

wa:l nama yinar-Ø banaga-ni
NEG that woman-**ABS** run-NFUT
'The woman didn't run.'

Let's see another example of a system strategy across world languages, this time from the remote Australian languages. Instead of an accusative-nominative alignment, these languages use an ergative-absolutive alignment. If you are familiar with German/Turkish/Latin terminology, these are cases just like Nominative and Accusative. However, their system is different, in that the ABS is found as alone argument in intransitive construction and as one of the two argument in a transitive construction, serving as the argument undergoing the action.

Introduction: How do we compare constructions within and across languages?

System strategy: the strategy is defined in terms **of the similarities and differences in the forms** of different constructions. (SYNCHRONY)



Blue = NOM and ERG strategies, light rose = ACC and ABS strategies (Papadimitriou et alii 2021)

Introduction: How do we compare constructions within and across languages?

Recruitment strategy: over time, the strategy **has been extended beyond its original construction**, covering the form of other constructions.

"There is a sense in which all strategies are probably ultimately **recruitment strategies** diachronically." (Croft 2022:19)

Example: in French, the **have possessive strategy** has been extended in other constructions beyond its original **possessive** construction, covering **physiological sensation predication**.

French

J' **ai** une petite tortue.

"I have a little turtle."

J' **ai** faim/peur/soif.

I have hungry/fear/thirst

"I am hungry/scared/thirsty".

Introduction: How do we analyze the structure of sentences in a particular language?

Comparative concept: Predicate nominal

Semantic class	**Propositional act**		
	reference	***modification***	***predication***
object	nominal phrase *head:* noun	possessive modifier/ genitive phrase	predicate nominal
property	property referring phrase	adjectival phrase *head:* adjective	predicate adjectival
action	complement (clause)	relative clause	verbal clause *head:* verb
(all semantic classes)	referring/argument phrase *head:* referent expression	attributive phrase *head:* modifier	clause: *head:* predicate

We are here, right?

We meet another American native language, Kiowa, from the small family of Kiowa-Tanoan.

Constructional Analysis: Predicate nominal, strategy: verbal copula

Kiowa (Kiowa-Tanoan, Stassen 1997:94)
té: kɔ́ýgú **bà-dɔ̄:**
all Kiowa.INV **2PL-be**
'You are all Kiowa'.

French (Indo-European, Croft 2022:18)
Elle **est** médécin
She **be.3SG.PRS** doctor
'She is a doctor'.

Do we have the same strategy in French and Kiowa? And what about English, Italian, … ?

And another language from our sample, Burmese, a language belonging to the big Sino-Tibetan family. This is the family of the Mandarin language.

Constructional Analysis: Predicate nominal, strategy: zero strategy

Burmese (Sino-Tibetan, Stassen 1997:90)
thu siʔthà
he soldier
'He is a soldier.'

Do we have the same strategy in French and Kiowa? And what about English, Italian, … ?

Classic Nahuatl was a language spoken in Mexico until the 16th century, where it was described by a Spanish missionary in the codex Florentine, which you can see in the picture. The codex also contains a quite long text in Classic Nahuatl. Nowadays, modern varieties of Nahuatl are still spoken in the area.

Constructional Analysis: Predicate nominal, strategy: verbal strategy

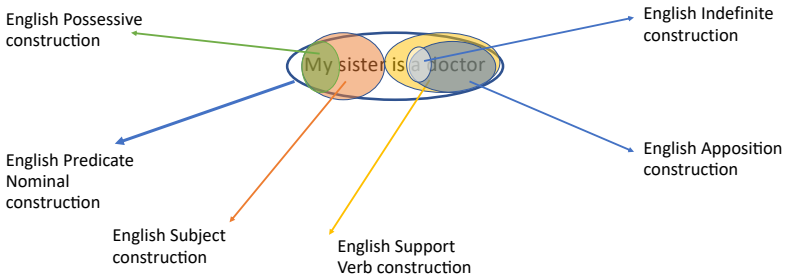
Classic Nahuatl (Uto-Aztecan)
ni-cihuātl
1SG-woman
'I am a woman.'

Cfr. with an instance of Classic Nahuatl Verb
ni-cuica
1SG-sing
'I sing'.

Classic Nahuatl uses verbal inflection.

How do we analyze the structure of sentences in a particular language?

More than often a construction is made by several constructions.



Interlinear Glossed Texts (IGT): How to read (and write) linguistic data

- **Language (Family, Source of the linguistic data)**
- **first line: original text** (**L1**)**;**
- **(optional) transliteration or phonetic text (L1)**
- **second line: interlinear morphemic glosses,** i.e**. grammatical** (in **Leipzig glosses**) and **lexical** (in **L2**) **morphemes; (aka IMT)**
- **third line: translation** in **L2.**

We use Leipzig glosses as the standard.

A common practice in Linguistics is to present linguistic data with glosses and translation, or the Interlinear Glossed Text. Linguistic data is then organised on three lines: the original text, the interlinear morphemic glosses and a translation. The original text is of course given in the original language (L1), while translation and lexical morphemes are given in an auxiliary language, L2, which is usually English. Grammatical morphemes are described according to the 'List of Standard Abbreviations' described in the Leipzig Glossing Rules.
Colour are for explanation only !

Interlinear Glossed Texts (IGT): How to read (and write) linguistic data

Modern Greek (J. K. Rowling, *Harry Potter and the Chamber of Secrets*, Greek trans. by K. Oikonomou)

θα σε πάρω μέσα στις αναμνήσεις μου τη νύχτα εκείνη

tha se páro mesa stis anamnίseis mou ti nýchta ekeíni

FUT you.ACC take into in.the memories my the.F.SG night.(F).SG that.F.SG

που τον έπιασ-α

pou ton épias-a

where him catch-AOR.1SG

'I'll take you into my memories of the night I caught him.'

In this case we have a mixed phonetic transliteration, using Latin scripts to represent Greek pronunciation and/or Greek scripts…

Interlinear Glossed Texts (IGT): How to read (and write) linguistic data

OMG, what all these punctuations and acronyms mean???

- Acronyms: refer to the Leipzig glosses;
- Quite often, you also need to describe the **morphological level**: split the word into morphemes by using the dash (-): έπιασ-α/*épias-a,* where –a is the morpheme and epias- the stem.
- A full-stop (.) between the gloss(es) and/or the acronym means that the glossed element (word or morpheme) bears different grammatical features (**cumulation**). For instance, Greek εκείνη (ekeini) is a feminine (F) singular (SG) meaning that: that.F.SG and –a is the first person singular (1SG) of the aorist (AOR).

Things are in fact more complicated, but here we give a simplified version of the Leipzig glosses.
Where, actually it should be e-pias-a, as the e- is also a morpheme indicating the aorist

Interlinear Glossed Texts (IGT): How to read (and write) linguistic data

yä-pitär	yäfəqr-u-n	zäfän	azzäfafän	← **object language**
GEN-Peter	love-DEF-ACC	song	sing:NML	← **IMT**

(1) Irish

Tá sí ag crú na mbó.

tɑ:	ʃi:	ə	kru:	nə	mo:
be(AUX)[ipfv]	PRO(3SG.F)[nom]	at	milk[ger]	DEF.ART[GEN.PL]	GEN.PL\cow

Interlinear Glossed Texts (IGT): How to read (and write) linguistic data

yä-pitär	yäfəqr-u-n	zäfän	azzäfafän	← **object language**
GEN-Peter	love-DEF-ACC	song	sing:NML	← **IMT**
'Peter's singing [i.e. his way of singing] the love song'				← **(free) translation**

(1) Irish

Tá sí ag crú na mbó.

tɑː	ʃiː	ə	kruː	nə	moː
be(AUX)[ipfv]	PRO(3SG.F)[nom]	at	milk[ger]	DEF.ART[GEN.PL]	GEN.PL\cow

'She is milking the cows.' (Christian Brothers 1993: 129)

Interlinear Glossed Texts (IGT): How to read (and write) linguistic data

(3) German

a. Der Schüler schreibt einen Brief.

deː-ʁ	ʃyːlʁ	ʃʁaɪp-t	aɪn-ən	bʁiːf
	S	V		O

'The student writes a letter.'

INDF = INDEFINITE REFERENCE.

ACC = accusative
ART= article
DEF = definite
NOM = nominative
3SG = third singular

Interlinear Glossed Texts (IGT): How to read (and write) linguistic data

(3) German

a. Der Schüler schreibt einen Brief.

deː-ɐ	ʃyːlɐ	ʃʁaɪp-t	aɪn-ən	bʁiːf
DEF.ART-NOM	student	write-3SG	INDF.ART-ACC	letter
	S	V		O

'The student writes a letter.'

INDF = INDEFINITE REFERENCE.

ACC = accusative
ART= article
DEF = definite
NOM = nominative
3SG = third singular

Chapter 2: *Propositional Act Constructions: The Skeleton of a Sentence*

Propositional acts: semantic classes and information packaging

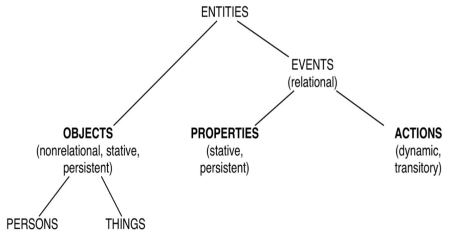
Semantic class	Propositional act		
	reference	***modification***	***predication***
object	nominal phrase *head:* noun	possessive modifier/ genitive phrase	predicate nominal
property	property referring phrase	adjectival phrase *head:* adjective	predicate adjectival
action	complement (clause)	relative clause	verbal clause *head:* verb
(all semantic classes)	referring/argument phrase *head:* referent expression	attributive phrase *head:* modifier	clause: *head:* predicate

Let's focus again on the functional part. What are these semantic classes and information packaging in more detail?

Propositional acts: semantic classes and information packaging

Semantic classes

- **Object**:
  - Non-relational: they exist in themselves; they do not need to make relation to other concepts
  - Static: they do not involve change over time
  - Stable (persistent): their identity lasts for their lifetime
- **Action**:
  - relational: they need someone or something to be performed;
  - Non-static (dynamic): they involve change over time
  - Non-stable (transitory): they can be started or stopped; they last only for a given amount of time
- **Property**:
  - Relational: they need someone or something to modify;
  - Static: they do not involve change over time
  - Stable (persistent): their identity (usually) lasts for their lifetime



Croft 2022:31

Ontology (= discourse about being) of semantic classes. We can define our semantic classes according to three binary features (yes/no)

Action and Object

Using other semantic (= cross-linguistic) features, we have the following more fine-grained classification:

- Action: (Vendlerian classes - *Aktionsart*, Vendler 1957)
  - **state** (+static, -stable, -telic): *know, be cool*
  - **activity** (-static, -stable, -telic): *run, dream*
  - **achievement** (-static, -stable, +telic): *run a mile, build a house*
  - **accomplishment** (-static, +stable, +telic): *win, be born*
- Object: (Rijkhoff 2002: 50-66 - *Seinsart*)
  - **mass** (-shape, +homogeneity): *water, money*
  - **count** (+shape, -homogeneity): *umbrella, bicycle*
  - **collective** (+shape, +homogeneity): *family, flock*

(this part is not on the textbook)
We have seen the PROTOTYPICAL semantic classes, but there is more to say, in fact.
An important classification for action is Vendler 1957, which is based on binary features (+/-).
A similar, less known classification is proposed for object by Rijkhoff and is based again on binary features: shape and homogeneity (and other features, but let's focus on these two for the moment).
Since this classification is based on semantics, we can still expect its cross-linguistic validity.

Action: *Aktionsart*

Three binary features:

- **static**: if +static, the action doesn't support strategies like English -ing. **I am being cool* (**state**) but *I am running in the park* (activity)
- **stable**: if +stable, the action doesn't support strategies like English *for*. **I was born for one day* (**accomplishment**) but *I have been building a house for three years* (**achievement**) and *I am running for three hours* (activity)
- **telic (also bounded)**: from Greek τέλος, 'goal'. Does the action have a goal and/or a limit? *I am running in the park* (activity: without a goal) vs. *I am running a mile* (**achievement**: the goal)

(This part is not on the textbook)
We already know the first two features, but let's add a test. Then, we should add a third feature, the telicity i.e., whether the action has a goal and/or culminates in something and/or has a bound/limit.
You can learn more on Aktionsart in Velupillai's handbook, pp.208-209

Object: *Seinsart*

- Two binary features (at least for our tripartite classification):
  - **shape**: it is the object counterpart of action **telicity**. Does the object occupy a fixed space i.e., has **boundaries**? Fluids like *water*, substances like *sand* or abstract concept like *money* do not. Object with +shape are compatible with numeral modification strategies such as *Two umbrellas* (**count**) and *Two families* (**collective**) but **Two waters* (**mass**)
  - **homogeneity**: is the object internally **homogeneous** i.e., is it made by **indistinguishable** parts like *drops of water* or *grains of sands*? If the object has +homogeneity, then it's compatible with degree modifier construction such has *You have more money than me* (**mass**) but **You have more umbrellas than me* (**count**)

Shape is modelled after action telicity. Like actions (run a mile, build a house), objects can be limited, occupying a fixed space. This objects are compatible with counting strategies.

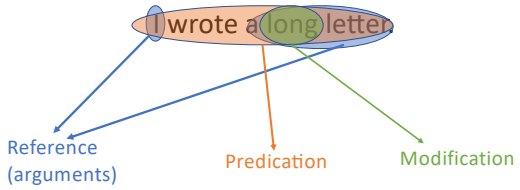
Propositional acts: semantic classes and information packaging

Three propositional acts: (file metaphor)

- **reference**: we open a new file (**referent**) or access an old file (**referent**) in the mind of the hearer;
- **predication:** we add information to this file, by asserting something. It is customary to refer to **referent(s)** in predication as **argument(s)**.
- **modification:** we enrich the file with **secondary information**, by modifying this referent with something.

Ok, let's go back to our textbook. We have the three propositional acts already seen: let's try to understand them better by using the file metaphor.
You probably have already guessed that the two essential propositional acts are reference and predication, while modification can be optional.

Propositional acts: semantic classes and information packaging



Again, we are working with parts of constructions, so a modification is embedded into a reference and references are embedded into a predication.

The major propositional act constructions and their structure

Constructions are often complex: they are made of **elements**, which are themselves **constructions**.

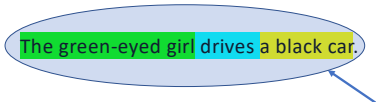
Two types of **elements**:

- **Head**: in a construction, the head is "**the most contentful word that most closely denotes the same function as the phrase (or clause) as a whole" (Croft 2022:33)**
- **Dependents:** in a construction, the **dependents** are simply everything else that are not the head and have a functional relation to the head.

As several of the previous examples show, constructions are often made by elements, themselves constructions. We can distinguish two elements in a construction: the head and the modifier. As always, we try to define these elements on a functional basis. In fact, we just have to define what the head is, as dependents are just elements that are not the head. So, the head is the word with the most important function in the construction; its function is so important that gives the function to the entire construction! We are going to see in the next slides what phrases and clauses: for the moment, just keep in mind that they are types of constructions.

The major propositional act constructions and their structure

Remember that constructions are often **nested/embedded within each other**: an element can be the **head** of a construction (cx) AND the **dependent** of another construction (cx).



Whole sentence (cx: sentence): drives (head), everything else (dependents)
The green-eyed girl (cx: phrase): girl (head), the / green-eyed (dependents)
green-eyed (cx: compound): headless
a black car (cx: phrase): car (head), a / black (dependents)

This is a phrasal analysis

In this example, girl and a car are both elements of the whole sentence AND heads of their phrase. In order to understand the embedding of constructions we have to split the sentence into phrases, a type of a construction. We have another type of cx here, a compound: what is its head?

The major propositional act constructions and their structure

Try to guess what is the head and what the dependent in the following constructions:

Cx: Compound
English
Taxi driver

Cx: sentence
Turkish

b. Evlerimizden gelmiyordum.
ev-ler-im-iz-ˈden ˈgel-mi-jor-d-um
house-PL-1SG.POSS-PL.POSS-ABL come-NEG-PROG-PST-1SG
'I was not coming from our houses.'

Cx: sentence
German
Ich weiss, dass er zu viel Arbeit hat.
I know that he too much work has
'I know that he has too much work'

Cx: phrase
French
Un chef-d'œuvre de patience.
A masterpiece of patience
'A masterpiece of patience.'

The major propositional act constructions and their structure

Three types of constructions:

- phrases:
  - referring phrase: a phrase whose function combines a referent and the reference propositional act e.g., English *the green eyed girl*; Turkish *güzel evlerimizden* 'from our beautiful houses"
  - attributive phrase: a phrase whose function combines a modifier and the modification propositional act e.g., [a hot] summer
  - admodification phrase: attributive phrase itself modified by an admodifier e.g., English [*a very hot*] *summer*; Turkish *[çok güzel] evlerimizden* '[from our very beautiful] houses'
- clauses:
  - predication clause: a clause whose function combines referring phrase(s) (here: argument phrases) with the predication propositional act through a predicate: English *'The green-eyed girl drives a black car'*; Turkish *ben evlerimizden gelmiyordum*
- complex predicates: a construction expressing predication, but with multiple morphosyntactic elements – sometimes discontinuous - instead of one: English *The soldier quickly walked off, vs. The soldier slowly sneaked in*

With respect to their size, we have two types of constructions: phrases and clauses. A referring phrase is a construction (cx) whose function is reference and a concept packaged as a referent, while a modifying phrase is a cx whose function is modification and a concept packages as a modifier.

If a modifier is itself modified by an admodifier, such as a degree like English very or Turkish cok, we have an attributive phrase. We move then to clauses, for instance predication clauses, in which we find an element functioning as a predicate that predicates one or more referring phrase, here called referring phrases. I said 'for instance' as predication clauses are just the most common type of clauses, but there are also other types of clauses which we are not going to discuss here.

Finally, we have predicates whose function is distributed across different morpho-syntactic elements. In the English example, the manner of motion is expressed by walked, the rate by quickly and the direction by off. Cfr. with the other example.

Noun, verb and adjective as comparative concepts: prototypical constructions

Semantic class	**Propositional act**		
	reference	***modification***	***predication***
object	nominal phrase *head:* noun	possessive modifier/ genitive phrase	predicate nominal
property	property referring phrase	adjectival phrase *head:* adjective	predicate adjectival
action	complement (clause)	relative clause	verbal clause *head:* verb
(all semantic classes)	referring/argument phrase *head:* referent expression	attributive phrase *head:* modifier	clause: *head:* predicate

Adjective, noun and verb can be taken as prototypical constructions i.e., things that are valid cross-linguistically. (I use yellow for noun, green for adjective and light blue for verbs). You also remember that we are on the what-side of the construction and we define it using semantic classes and information packaging.

OK: but why these constructions and not, say, property referring phrase or predicate nominal? Primarily, because of their cross-linguistic frequency in usage: they are the most three common constructions in languages, the richest in expressing related features (for instance, if a language has gender then it will probably use gender first for noun and then for other word classes) and have the shortest encoding strategy (for instance, the strategy to express a complement clause is no longer than the strategy expressing verb). We will come back to this topic later.

Noun, verb and adjective as comparative concepts: prototypical constructions

Typological (cross-linguistic adequate) definitions:

- noun: the head of a phrase that refers to object(s) (referring phrase);
- adjective: the head of a phrase that modifies by property(ies) (attributive phrase);
- verb: the head of a phrase that predicates of action(s) (clause).

If we were using morpho-syntactic strategies like 'a noun inflects for case', 'an adjective agrees for gender' or 'a verb inflects for tense', we wouldn't be able to formulate definitions valid for all languages.

For instance, only Greek, German and Turkish have cases in our sample; English adjectives does not agree for gender (and number). The tense criterium seems to work for all languages of our sample, but there are many languages that do not morphologically inflect for tense, for instance Mandarin Chinese and Vietnamese.

These definitions work in order to define the strategies that encode the three comparative concepts within a single languages: we can speak of language-specific word classes vs. prototypical (universal parts-of-speech).

Noun, verb and adjective as comparative concepts: prototypical constructions

- Is 'doctor' a noun in I am a doctor? No, it looks like a noun as in English is encoded like a noun (plus a copula)
- Is 'beautiful' an adjective in 'the fact of being beautiful is important on social networks'? No, it looks like an adjective, as in English is encoded like an adjective (plus 'the fact of COPULA –ing')
- Is 'playing' a verb in 'I like girls playing guitar'? No, it looks like a verb as in English is encoded like a verb (inflected in the gerundive form)

As **comparative concepts** these are: predicate nominal, property referring phrase and relative clause, respectively.

As **English-specific constructions**, these are: English Predicate Nominal, English Property Referring Phrase and English Relative Clause.

Does it sound counter-intuitive? Let's make a cross-linguistic exercise.

English
I am not a doctor
1SG be.COP NEG indf doctor

Classical Nahuatl
ah-ni-tīcitl
NEG-1SG-doctor
'I am not a doctor'.

So, in Classical Nahuatl nouns are verbs? Not really, English and Classic Nahuatl simply use two different strategies to express the predicate nominal: a copula, which is also used in English for predicate adjectival, and a verbal inflections, which is also used in Nahuatl for verbs. Their head is doctor in English and ticil in Nahuatl, but these heads are not nouns either, they are head of predicate nominals. We can (in a cross-linguistic fashion) use noun only when we have a referring phrase whose head is an object, nothing more. Refer also to slide no. 49, in which we have contrasted the different strategies for adjective, nominal predicate and adjectival predicate in Turkish and in English.

More on the structure of propositional act constructions

- head of **referring phrase**: noun (**prototype**), referent expression (**general**). For instance, in German ***Wandern*** *macht Spaß* 'Hiking is fun', *Wandern* is not a noun, but still functions as a referent expression;
- head of **attributive phrase**: adjective (**prototype**), modifier (**general**). For instance, in Turkish ***elli*** *ağaç* '50 trees', *elli* '50' is not an adjective, but still functions as a modifier;
- head of **clause**: verb (**prototype**), predicate (**general**). For instance, Greek είναι **γιατρός** 'she is a doctor', γιατρός is not a verb, but functions as a predicate.

Let's distinguish between a prototypical and a general term for the head of referring & attributive phrase and clause.

Nonprototypical propositional act constructions

What about the other six constructions in Semantic class X Propositional Act (Information Packaging) matrix? These are non-prototypical constructions, in the sense that:

- they are less common than noun, adjective and verb, esp. in spoken discourse;
- they may use different morpho-syntactic strategies with respect to prototypes, but they can also **recruit strategies** from prototypes.

Here's a list:

- Property referring phrase construction (property reference);
- Possessive modifier construction (modification by object);
- Predicate nominal construction (predication of object);
- Predicate adjectival construction (predication of property);
- Complement clause construction (action reference);
- Relative clause construction (modification by action).

When speaking, it's more common to use just noun, adjective and verb rather than, say, property referring expression, action noun or even relative clauses…
In the following, we will try to contrast the first two constructions for languages of our sample; we have already contrasted the third and the fourth, and we are going to take a look to the fifth and sixth construction.

Complement clause construction

- This is a broad term encompassing both Nominalizations and Complements Clauses construction: **all denote reference to action**.
- Nominalization and Complemente Clauses are quite different in our European languages, but e.g., in Sino-Tibetan languages there is a continuum. In English we have: Nominalization, Gerund, Infinitival Complement and Finite Complement.
- Since clauses have arguments, complement clause constructions may encode these arguments. Compare 'The enemy destroys the city' vs. 'The destruction of the city by the enemy'.

In many languages of the world our nominalizations serve to create subordinate clauses, which corresponds to English Infinitival and Finite Complement.
And many languages offer the possibility to express clausal arguments in the complement clause construction, even in the short form of Nominalizations.

Relative clause construction

- Again, we have a broad term including different - at least for our languages – constructions expressing the more general function of **modification by action.**
- English constructions: Present, Past/Passive Participle, Infinitival Relative Clause, Finite Relative Clause.
- Argument encoding. Compare 'The gold was stolen by the Mafia' vs. 'The gold stolen by the Mafia'.

Three principles of the mapping between form and function

Three principles (Croft 2022:45-49)

1. **First principle of information packaging/construal**: any semantic content may be packaged in any way, in order to serve the joint goals of the interlocutors in discourse.
2. **Second principle of information packaging/construal**: the nature of reality, e.g. the semantic characteristics of semantic classes, favors (or disfavors) certain ways of packaging that information.
3. **Third principle of information packaging/construal*:*** the relationship between form and meaning—what sort of construction a word with a particular meaning occurs in— is a matter of cultural convention, that is, the linguistic conventions of the speech community.

The main reason behind such complexity in the structures of world languages lay in the two-fold nature of function: semantics and information packaging.
We have already seen that each semantic class can be packaged using the three propositional acts; we have also seen that the definition of constructions cannot relay only on semantic classes, like 'verbs are actions, nouns are objects and adjectives are properties'. However, the second principle tells us that this naïve assumption has some truth: it is more natural to refer to stable and relation things (object), predicate non-relational and transitory things (actions) and modify by using non-relational but stable things (properties).
Again, we can of course refer to non-relation things (property and action) and modify by transitory things (actions), but it is less natural.
We can start from the first principle (you can express everything) and see how the other two principle somewhat constrain it. The second one is a constrain from reality, while the third one is a cultural constrain.

Three principles of the mapping between form and function

Grammaticalization is a recurrent pattern in the structure of the world languages: **over time** (diachrony) **a grammatical construction acquires a new function**.

- Remember the **recruitment strategy**? "all constructional strategies are ultimately instances of recruitment" (Croft 2022:48)"
- The process is gradual, so grammatical constructions are often **a blend of old and new form (strategy) and function (semantics and information packaging)**.
- Speakers are **creative** (they always need new expressing functions) but are also **conservative** in terms of grammatical constructions, recycling old strategies for new function.

Due to the high variability of cultures, it is very difficult to make predictions on grammatical structures of the world languages. However, typologists and historical linguists have already been discovered several patterns that account for some variability in world language. One of these patterns is grammaticalization, or the idea that grammatical constructions constantly evolve, changing functions but often retaining old forms.

Let's take the expression 'You better leave now', in which 'better' functions just like modal auxiliary verbs i.e., You must/shall/may/ought to leave now. Why an adjective in a comparative form is found in this position? Because the expression is actually a reduction of You had better leave now, but eventually the original auxiliary dropped and 'better' acquired the position (and the function) of an auxiliary.

Recruitment strategies for nonprototypical constructions

semantic information packaging (IP) strategy: the non-prototypical construction *recruits* strategies from the prototypical construction with respect to the **semantic class.**

Semantic class	Propositional act		
	reference	*modification*	*predication*
object	**a** doctor		(I am not) **a** doctor
property			
action			

actual information packaging (IP) strategy: the non-prototypical construction *recruits* strategies from the prototypical construction with respect to the **propositional act**:

Semantic class	Propositional act		
	reference	*modification*	*predication*
object			**ah-ni-**tīcitl NEG-1SG-doctor
property			
action			**ah-ni-**chōco NEG-1SG-cry

Here, the English Nominal Predicate recruits the strategy from the English Noun (reference, the indefinite article): the morpho-syntax is taken by the prototype wrt to the semantic class; by contrast, in Classic Nahuatl, the morpho-syntax is taken from the Verb (verb inflections), as Classic Nahualt looks here at the information packaging (a Nominal predicate is an action) rather than the semantic class.

Recruitment strategies for nonprototypical constructions

overt coding strategy: 'something' (morpheme, word, tone) is added to the prototypical recruited strategy.

	reference	modification	predication
object	The girl dreams.		
property	The wise **one** dreams.		
action			

zero coding strategy: the prototypical recruited strategy is used as-is.

	reference	modification	predication
object	Το κορίτσι ονειρεύεται. Das Mädchen träumt		
property	Ο σοφός ονειρεύεται Der Weise träumt.		
action			

In Greek, German and Engglish, the strategy for property reference is recruited from the prototypical function with respect to the information packaging; English adds a 'one', so it's an overt coding strategy, while Greek and German doesn't add anything (zero coding strategy). Arrow represent the sense and direction of recruitment.

Recruitment strategy

- **hybrid information packaging (IP) strategy**: the non-prototypical constructions recruit strategy **from both actual and semantic information packaging (IP) strategy;**
- To be honest, it was not fully correct defining most of the the previous strategies as either **actual or semantic IP strategies;**
- in fact, English Predicate Nominal is an example of **hybrid strategies.**

Let's review together the previous construction

English Predicate Nominal

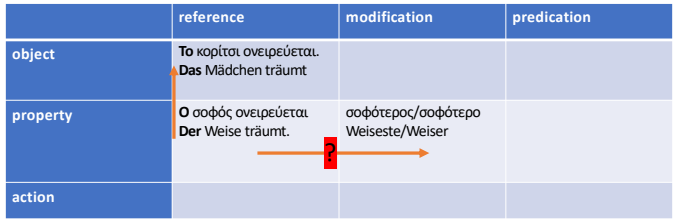
	reference	modification	predication
object	**A** doctor		I *am not* **a** doctor. I *won't be* a doctor. I *may be not* a doctor.
property			
action			She plays. She *will* play. She *may* play.

From reference of object = definite article (definiteness)
From predication of action = verbal categories (tense, mood)

Hybrid

In fact, English Predicate Nominal also recruits strategy from the predication of action, since it inflects for TENSE and MOOD.

Greek and German Property Referring Phrase



From reference of object = definite article (definiteness)
From modification of property = degree ???

Hybrid?

EL/DE Property references recruits the strategy from object reference (definite articles). Do they also recruit strategies from modification by property (semantic IP strategy)?

English Complement Clause (strategy = gerundive)

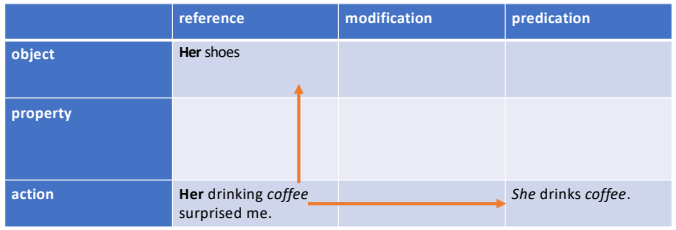
	reference	modification	predication
object	**Her** shoes		
property			
action	**Her** drinking *coffee* surprised me.		*She* drinks *coffee*.

From reference of object = possessive (her)
From predication of action = argument (she, coffee)

Hybrid

English Gerundive is a hybrid strategy, as it recruits an actual IP strategy from the object reference (the possessive) and semantic IP strategy from predication. (arguments)

Greek and German Complement Clause (strategy = gerundive)



From reference of object =
From predication of action =

Hybrid

What about German and Greek Gerundives?

Two crosslinguistic universal of grammatical strategies

- **structural coding**: "A lexical class used in a **nonprototypical propositional act function** will be coded with **at least as many morphemes as in its prototypical function.**"
- **behavioral potential**. A lexical class used in a **nonprototypical** propositional act function will also **have no more grammatical behavioral potential** than in its prototypical function.

Croft 2022 (55-56)

The first cross-linguistic universal is quite straightforward: non-prototypical classes have at least many morphemes as the respective prototypical classes. For instance, the predicate adjectival construction in English has more morphemes than the adjective (it has the copula), and so on.

Behavioral potential is more subtle: do you remember when we have tried to construct German or English Action Nominalizations in the past or with different mood? We have seen that it was quite difficult or even possible, and sometimes we had to resort to other strategies. This is an example of Behavioral Potential: English Action Nominalizations have no more grammatical categories than English Verbs or English Nouns (in this case, English Action Nominalization has less categories)

Two crosslinguistic universal of grammatical strategies

- **structural coding**: "A lexical class used in a **nonprototypical propositional act function** will be coded with **at least as many morphemes as in its prototypical function.**"

	Reference	**Modification**	**Predication**
Objects	*vehicle*	*vehicle**'s**, vehicul-**ar**, **of/in**/*etc. *the vehicle*	***be a** vehicle*
Properties	*white-**ness***	*white*	***be** white*
Actions	*destruc-**tion**, **to** destroy, destroy-**ing**, **that**...destroy*	*destroy-**ing**, destroy-**ed**, **which/that…**destroy*	*destroy*

On the diagonal we see the three prototypical classes, represented here by vehicle, white and destroy. When the semantic class is used with a different, non-prototypical propositional act function, it takes at least as many morphemes as in the prototype. In all cases, less.

Historical note: this is also known as markedness, which traces back to the Prague School (or Prague Linguistic Circle) of the 20s of the 20th century. This version of markedness, or 'typological markedness', differs from the older one as the non-prototypical form can be marked with at least as many morphemes, while the Prague School theory of markedness has that non-prototypical forms must have more morphemes than non prototypical. This is untenable, as we have seen for instance with German/Greek Adjectives used as Nouns without further marking (Property Referring Phrase).

Two crosslinguistic universal of grammatical strategies

- **behavioral potential**. A lexical class used in a **nonprototypical** propositional act function will also **have no more grammatical behavioral potential** than in its prototypical function.

  Typical behavioral potential for reference: number, gender, case, definiteness; indexation of possessor

  Typical behavioral potential for modification: degree (simple, equative, comparative, superlative); indexation of head noun in number, gender and case

  Typical behavioral potential for predication: tense, aspect, modality, polarity (TAMP); indexation of subject (and object) arguments in person, number and/or gender

And here's a list of grammatical features, again from Croft. This will be relevant in defining how prototypical and non-prototypical constructions behave across languages.