

Elements of theory of grammar

1. The unbounded character of natural languages.

(1) A major distinctive property of human language is that its scope is unbounded. We can constantly create and understand new sentences, sentences that we have never encountered in our previous experience with the language (rule-governed creativity). Moreover, there is no upper limit to the length and complexity of a sentence (there is nothing like “the longest English sentence”). How can it be that finite organisms with a finite experience can master an unlimited set of potential messages? This is the foundational question for modern formal linguistics and generative grammar.

(2) “The central fact to which any significant linguistic theory must address itself is this : a mature speaker can produce **a new sentence** of his language on the appropriate occasion, and other speakers can understand it immediately, though it is equally new to them. **Most of our linguistic experience, both as speakers and hearers, is with new sentences**; once we have mastered a language, the class of sentences with which we can operate ... is so vast that... we can regard it as infinite”
Chomsky Current Issues, 1964

(3) “The Basic Property of human language: each language yields a digitally **infinite array of hierarchically structured expressions** with systematic interpretations at interfaces with two other internal systems, the **sensorimotor system** for externalization and the **conceptual system** for inference, interpretation, planning, organization of action, and other elements of what is informally called «thought»”.

N. Chomsky & R. Berwick (2016) *Why only us – Language and Evolution*. MIT Press, p. 88

(4) René Descartes and syntactic creativity :

« Car on peut bien concevoir qu’une machine soit tellement faite qu’elle profère des paroles... mais non pas qu’elle **les arrange diversement**, pour répondre au sens de tout ce qui se dira en sa présence, ainsi que les hommes les plus hébétés peuvent faire » (p. 86)

« Or, par ces ...mêmes moyens, on peut aussi connaître la différence qui est entre les hommes et les bêtes. Car c’est une chose bien remarquable, qu’il n’y a point d’hommes si hébétés et si stupides, sans en excepter même les insensés, qu’ils ne soient capables **d’arranger ensemble diverses paroles, et d’en composer un discours par lequel ils fassent entendre leurs pensées** ; et qu’au contraire il n’y a point d’autre animal tant parfait et tant heureusement né qu’il puisse être, qui fasse le semblable.

(R. Descartes DM 1637)

(5) Galileo Galileo and the combinatorial character of the alphabetic writing system :

“...ma sopra tutte le invenzioni stupende, quale eminenza di mente fu quella di colui che s’immaginò di trovar modo di **comunicare i suoi più reconditi pensieri** a qualsivoglia altra persona, benché distante per lunghissimo intervallo di luogo e di tempo? Parlare con quelli che son nell’Indie, parlare a quelli che non sono ancora nati, né saranno se non di qua a mille e diecimila anni? E con qual facilità? **Con i vari accozzamenti di venti caratteruzzi** sopra una carta.” (p. 130)
(G. Galilei, Dialogo, 1630/1970)

(6) - Galileo refers to the «invention» of alphabetic writing, a cultural product.
 - But this «invention» is made possible by the discovery of the constitutive property of a natural object, the combinatorial character and hierarchical organization of natural languages.

(7) „[die Sprache] muss daher **von endlichen Mitteln einen unendlichen Gebrauch machen...**“
 “[language] must therefore make **an infinite use of finite means...**”
 W. von Humboldt (1836)

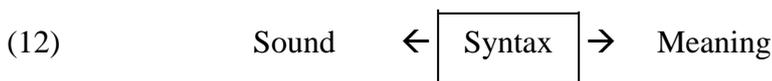
(8) Chomsky (1957) Syntactic Structures: knowing a language amounts to possessing a computational system, a generative grammar, which consists of

- a. a **lexicon**, and
- b. a **system of rules** which combine words to form higher order entities (phrases, sentences)

(9) the computational rules are **recursive**, i.e., they can reapply to their own output. Recursion is the fundamental formal property responsible for the unbounded nature of the system. Animal communication systems typically lack recursive mechanisms, so that animal communication systems typically are not unbounded. In parallel, many animal species have remarkable capacities of rough estimation of quantities, but generally lack the capacity of counting precisely.

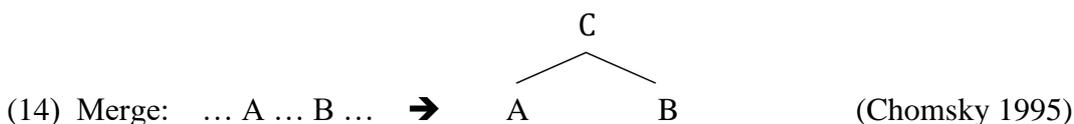
- (10)a Mary’s sister’s friend’s... car
- b La macchina ... dell’amico della sorella di Maria
- c John thought [that Mary had said [that her husband believed [that]]]

(11) Language is sound with meaning. The computing machine must be able to generate pairs of sounds and meanings over an unbounded domain



Interface representations of sounds are accessible to sensori-motor systems of sound articulation and perception (or articulation and perception of signs in sign language). Interface representations of meaning are accessible to conceptual-intentional systems: systems of concepts, systems of communicative intentions in our interactions with others.

(13) About half a century of debate on the nature of recursion in natural language led to the conclusion that there is a fundamental recursive rule, which is about the simplest combinatorial rule one can imagine:



In words, “take two elements A and B and put them together to form a third element C”. The rule is recursive because it can reapply to its own output, e.g., C can be merged with another element D and so on, indefinitely

(15) A minimal grammar for natural languages should contain at least two mechanisms:

- i. A **structure-building** mechanism (Merge)
- ii. A mechanism expressing **dependencies** between positions

A kind of dependency that has been studied in detail in the last 50 years or so is **movement**: an expression is often interpreted in a position different from the position in which it is pronounced

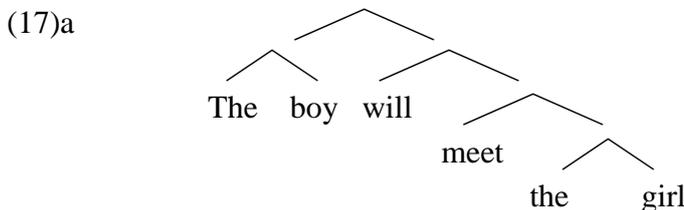
- (16)a We should read [this book]
b [Which book] should we read ___?
c [Which book] do you think [we should read ___]?
d [Which book] do you thin [the professor said [we should read ___]]?
.....

(16)b' We should read [which book]



2. The hierarchical nature of syntactic representations. (Lingua 130, Introduction)

Words are combined by merge in ways which give rise to hierarchically organized representations:



(17)b [[the boy] [will [meet [the girl]]]]

The structural organization expressed by (15), of course, is not an *a priori* truth: it represents an empirical hypothesis on the organization of the sentence, which can be tested by various means:

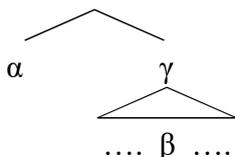
- by showing that some sequences of words behave as a unit with respect to certain syntactic manipulations, e.g., they can be displaced or elided together (structure dependence of rules)
- by studying the intonational contour of the sequence (Selkirk 2011, Bocci 2013);
- through various kinds of psycholinguistic experimentation (e.g., in the time-honored tradition of click location: Cohen & Mehler 1996 and references cited there);
- through brain imaging techniques (e.g., of the kinds used in Musso et al 2003, (see also Moro, The Boundaries of Babel) who observe differential activation in language areas for hierarchical and non-hierarchical dependencies, or in Pallier, Devauchelle, Dehaene 2011, who identify brain areas sensitive to the depth of the tree hierarchy, as opposed to linear length), etc.

3. C-command

This is a fundamental hierarchical relation, which has pervasive consequences in natural language syntax and semantics.

Let us say, as a first approximation, that a node α c-commands a node β in the following configuration, abstracting away from linear order:

(22) α c-commands β in the following configuration:



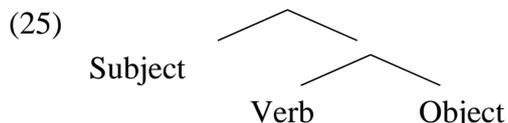
(23) α c-commands β iff β is dominated by the sister node of α . (Reinhart 1976, Chomsky 1995)

Two sister nodes are those which are combined by merge.

C-command plays a key role in referential dependencies:

- (24)a John criticized himself
b *Himself criticized John

This follows from binding principles capitalizing on the hierarchical relation of c-command. If, as we must assume for a number of independent reasons, the hierarchical structure of the sentence is roughly the following:



and we define binding in terms of c-command, we get the right result: subjects c-command objects, hence can bind them, but not vice-versa.

This analysis immediately generalizes to other kinds of structural relations: e.g., the possessive can bind the complement of the noun

- (26)a [John's picture of himself] was published in the newspaper
b *[Himself's picture of John] was published in the newspaper

Moreover, the system reduces potential ambiguities: in

(27) [John's brother] [criticized himself]

the anaphor *himself* can only refer to John's brother, not to John.

That the relevant property depends on hierarchy, not on linear order is shown by the fact that, when hierarchy and (surface) order don't match, hierarchical considerations prevail

- (28)a Maria si presentò
 b Si presentò Maria
 c La sorella di Maria si presentò ‘Maria’s sister introduced herself’

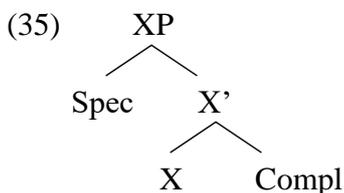
so, precedence is neither a necessary, nor a sufficient condition for anaphor binding.

3. Classical X-bar Theory (Chomsky 1970, Jackendoff 1977)

Two fundamental ideas:

- (33)a. Each phrase is the projection of a head. (Harris 1949)
 b. Phrases have a fundamentally uniform structure,

- (34) NP (I saw) [many pictures of Mary]
 (Ho visto) [molte foto di Maria]
- VP (I have) [always excluded this hypothesis]
 (Ho) [sempre escluso questa ipotesi]
- AP (I am) [very proud of Mary]
 (Sono) [molto fiero di Maria]
- PP (I’ll be back) [right after the vacation]
 (Tornerò) [subito dopo le vacanze]



4. An argument for the structural asymmetry specifier-complement

- (36)a [John_i’s brother]_k saw himself_{k, *i} in the mirror
 b [Il fratello di Gianni_i]_k ha visto se stesso_{k, *i} nello specchio

(37) Principle A of Binding Theory: an anaphor must be bound by an antecedent which c-commands it in a local domain.

- (38)a [John’s [picture of himself]] was published in the newspaper
 b *[Himself’s [picture of John]] was published in the newspaper

- (39)a [my friends’ [pictures of each other]] were published
 b * [each other’s [pictures of my friends]] were published

5. Ordering parameters

- (40) John loves Mary

S V O

(41) ...dass Hans Maria liebt (German)
 that H M loves
 S O V

(42) [V O] [O V]

French : Jean [aime Marie]

Latin : Tullius [Flaviam amat]

Chicewa: Njuchi [zi-na-wa-lum-a alenje]
 bees bit hunters

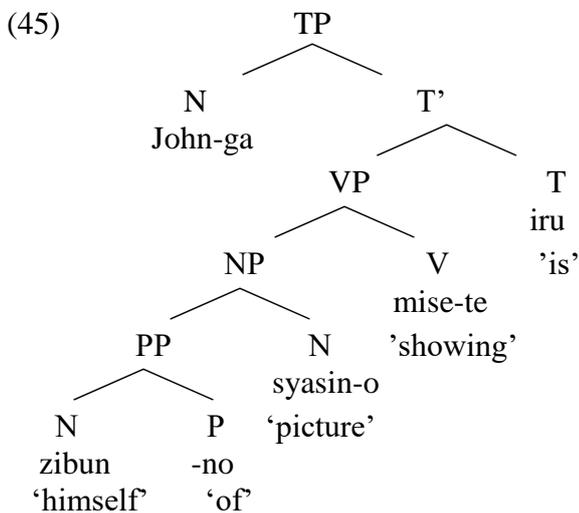
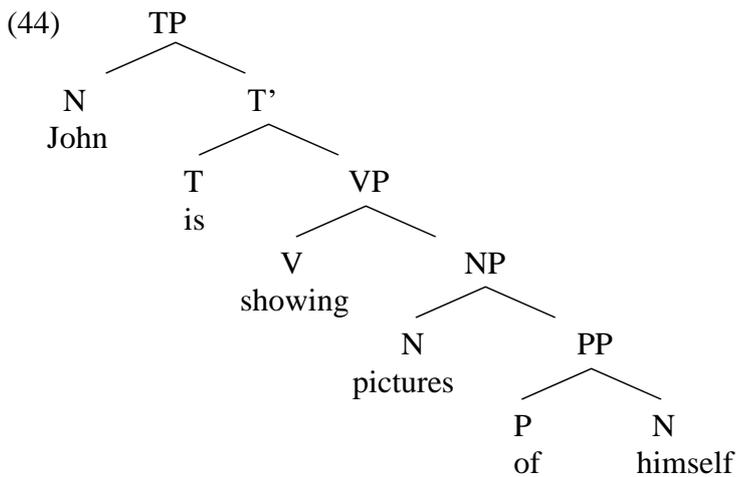
Japanese: John-ga [Mary-o butta]
 John Mary hit

Edo: Ozo [mien Adesuwa]
 Ozo found Adesuwa

Navajo: Ashkii [at'éeéd yiyiiltsa]
 Boy girl saw

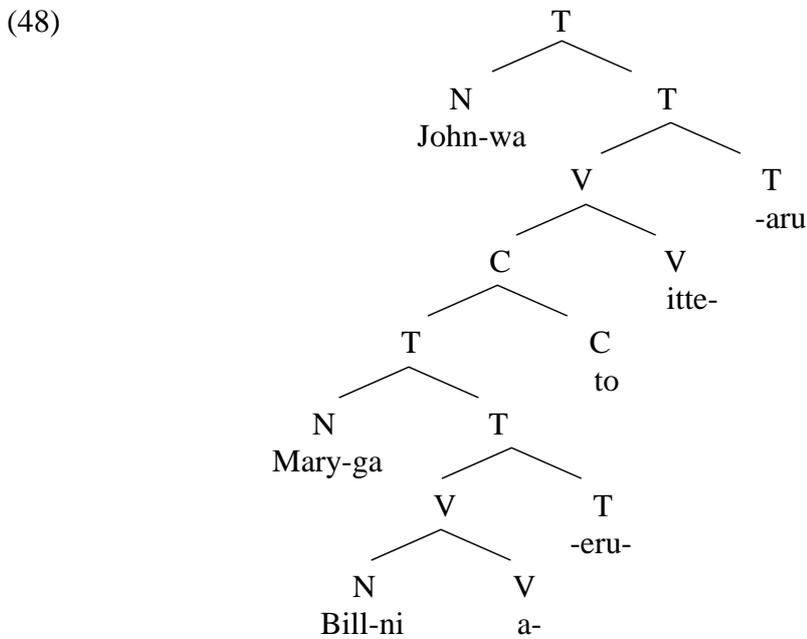
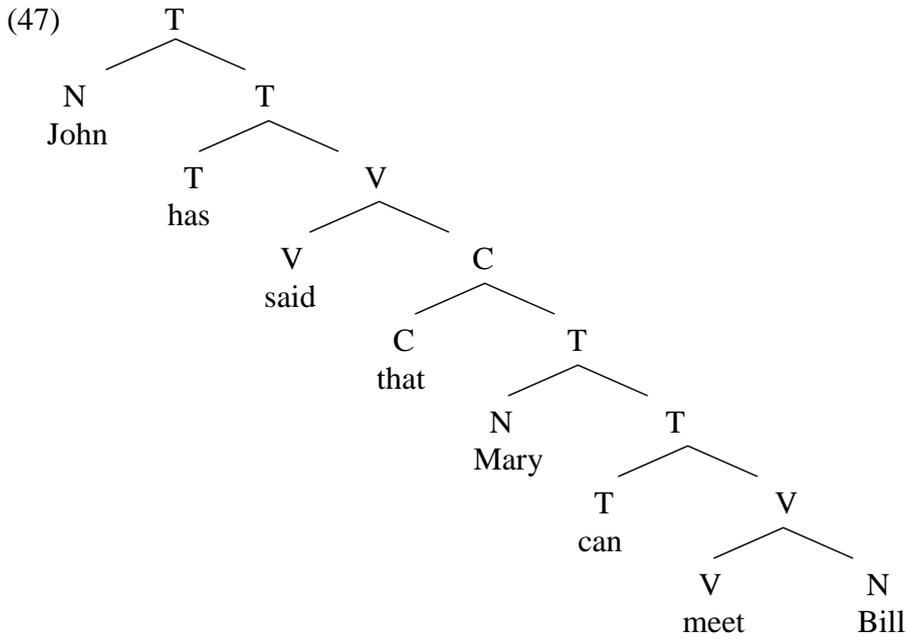
(43)a V precedes/follows O

b Head precedes/follows complement



(46)a John has said [that Mary can meet Bill]

b John-wa [Mary-ga Bill-ni a -eru- to] itte-aru.
 'John-Top [Mary-Nom Bill-Dat meet - can - that] said - has'



6. Structures and typological generalisations

(46) Greenberg (1963): VO languages tend to be prepositional, OV languages tend to be postpositional:

VO : 19	OV : 11
PO : 16	OP : 11
OP : 3	PO : 0

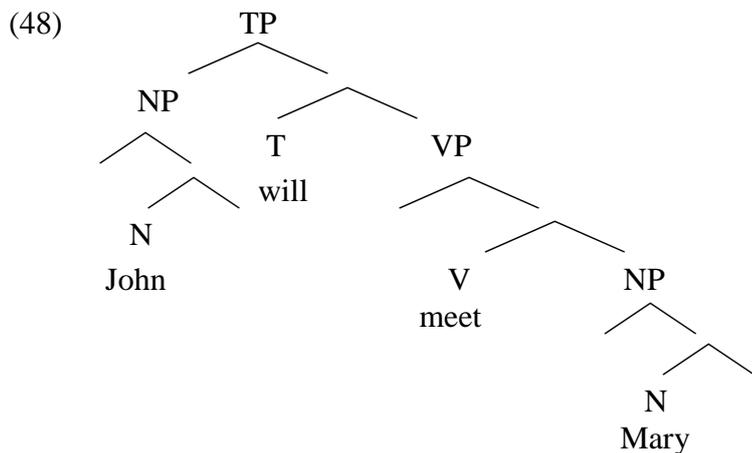
(47) Greenberg (1963): VO languages tend to have Aux before V, OV languages tend to have Aux after V:

VO: 11	OV: 8
Aux V: 10	V Aux: 8
V Aux: 1	Aux V: 0

Greenberg's generalizations hold statistically: languages tend to fix ordering parameters harmonically (with the same value for all categories), but disharmonic fixations are possible.

7. The head of the sentence.

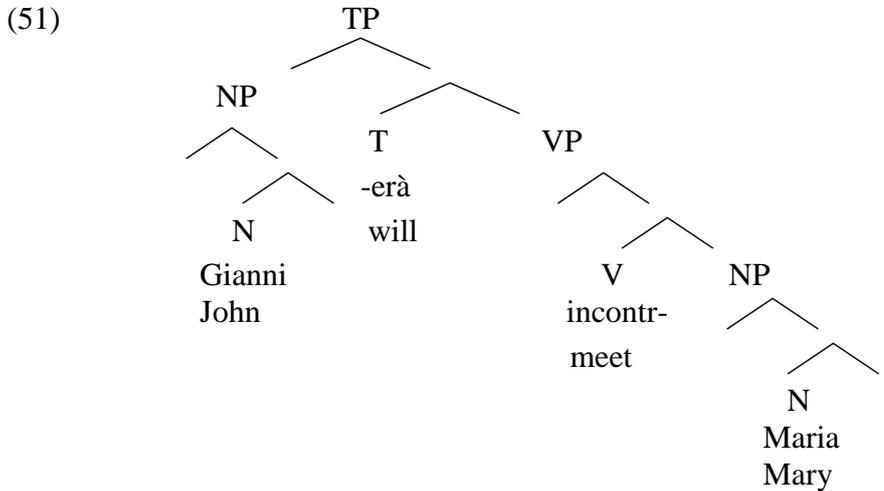
If all categories are headed, what is the head of the sentence? In VO languages the element marking tense (when it is an independent word) typically occurs between the subject and the predicate:



- (49)a Im nuo dat (Jamaican: Durrleman 2001)
 'He knows that'
- b Im en nuo dat
 'He PAST know that'
- c Im wi nuo dat
 'He FUT know that'

8. V to T movement.

- (50)a John will meet Mary
 b Gianni incontrerà Maria
 c Gianni incontr+erà Maria



(52) Vata (Kru language, spoken in Côte d'Ivoire : Koopman 1983):

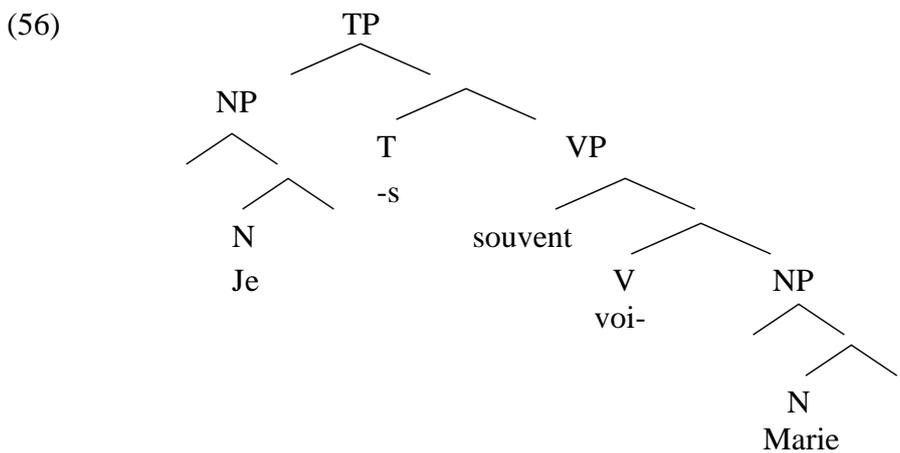
- a. à lî saká
 'we eat rice'
- b. à lā saká lî
 'we have rice eaten'

with perfective aspect, T is realized as Aux ; with durative aspects T is not realized as an autonomous word. In this case, V moves to T, which determines the SVO order (Koopman 1983).

- (53)a J' ai souvent vu Marie (French : Pollock 1989, Emonds 1978)
 b I have often seen Mary

- (54)a * Je souvent vois Marie
 b I often see Mary

- (55)a Je vois souvent Marie
 b * I see often Mary



- (57)a John will [often meet Mary]
 b Jean rencontr-era [souvent _____ Marie]
 c Gianni incontr-erà [spesso _____ Maria]

- (38)a Jean (n') a pas mangé la soupe
 John has not eaten the soup
 Gianni non ha più mangiato la zuppa

- b Jean (ne) mange pas la soupe
 *John eats not the soup
 Gianni non mangia più la zuppa

- c * Jean (ne) pas mange la soupe
 John does not eat the soup
 * Gianni non più mangia la zuppa

- d Pour (ne) pas manger la soupe
 In order to not eat the soup (Pollock 1989)
 *Per non più mangiare la zuppa (Belletti 1990)

- (59) V to T It Fr Engl
 +fin + + -
 -fin + - -

NB : we don't have case IV : finite paradigms are generally richer.

- (60) Jean ne voit pas souvent Marie
 John does not often see Mary

7. Diachrony.

(61) Roberts (1993) "[In English] V-movement began to decline in the latter part of xvi century and was lost from colloquial language in the xvii century":

- a If I gave not this accompt to you (1557)
 b In doleful way they ended both their days (1589)
 c I know *not*...
 d go *not* to Wittenberg
 e I speak *not* to him (Hamlet)

(62)

- Is: cast Ip: cast(-e)
 IIS: cast-est IIp: cast(-e)
 IIIS: cast-eth IIIp: cast(-e)

(63) There seems to be a relation between the richness of the verbal morphological paradigm and the movement of the lexical verb to the inflectional system.

(64) "Richness" is sometimes calculated on the basis of number of distinct infl morphemes in core paradigms (Roberts 1993, Vikner 1997, Rohrbacher 1999 (1st-2nd distinction in at least one number of 1 tense)), sometimes in terms of the possibility of cooccurrence of T and Agr morphemes (not in Modern English: call-s, call-ed, *call-ed-s, Bobaljik 2001).

8. Modern English –s

One case which remains problematic is the fact that in Modern English the residual inflectional system –s/-ed appears on the lexical verb which has remained VP-internal.

(65) John often sees Mary

(66) John -s often see Mary →

John ___ often see-s Mary

Chomsky (1957): in English T hops to V (affix hopping), but the process is blocked by certain intervening elements like negation; in this case, the auxiliary do is inserted to preserve morphological well-formedness:

- (67)a He -s see Mary
b He -s not see Mary
c -s he see Mary?

9. Microcomparison of the Scandinavian languages

(68) Icelandic (heyra ‘hear’)

present: heyr-i, heyr-ir, heyr-ir, heyr-um, heyr-ið, heyr-a

preterite: heyr-ði, heyr-ði-r, heyr-ði, heyr-ðu-m, heyr-ðu-ð, heyr-ðu

(69) Danish (høre ‘hear’)

present: hør-er

preterite: hør-te

(70) ... að hann keypti ekki [___ bokina] (Icelandic)
‘ that he bought not the book’



(71) ... at han ikke [købte bogen] (Danish)
‘ that he not bought the book’

NB : Faroese has a verbal paradigm of intermediate richness between Icelandic and Continental Scandinavian, and verb movement to T is unstable in the system (one dialect has it, another does not)

10. VSO

(72)a Cheannaigh siad teach anuraidh (Irish)
‘Bought they a house last year’

b Chuala Roise go minic an t-amharan sin
‘Heard Roise often this song’

Are VSO languages « flat » languages, without a VP ? No: standard arguments for VP apply to these languages (the subject asymmetrically c-commands the object for binding, idiomatic expressions of « John kicked the bucket » type normally concern V and O, not V and S, etc.).

Emonds (1981) : VSO is derived from S V O via verb movement to a functional head :

(73) X [NP [V NP]]

(74)a Cana i yfory (Welsh)
 'will-sing I tomorrow'

b Bydda i 'n canu yfory
 'will-be I singing tomorrow'

11. The Complementizer system (CP)

Embedded clauses are typically introduced by **complementizers** like *that, if*, etc. In main clauses the C position may be used as landing site for movement of T in special constructions, like interrogatives:

(75)a [___ C [you will [meet who]]]
 b [Who will [you ___ [meet ___]]]

(76)a Tu as vu qui ?
 b Qui as [tu ___ vu ___]?

In embedded clauses, e.g. relative clauses, we often have a Doubly Filled C effect:

(77)a Here is the man [who C [I saw ___]]
 b Here is the man [___ that [I saw ___]]
 c Here is the man [___ C [I saw ___]]
 d *Here is the man [who that [I saw ___]]

but in many varieties a wh element and that can co-occur:

(78)
 Middle English: The book which that I read
 Belfast English : I wonder which book that he read
 Dutch dialects: Ik weet niet [wie *dat* [Jan ___ gezien heeft]]
 Québec: Qui que tu as vu?
 Veneziano: Cossa che la magna?
 Italian: Che bel libro che hai letto! (exclamatives)

(79) Belfast English (Henry 1994):
 a. I wonder which book that you read
 b. I wonder which book did you read
 c. *I wonder which book that did you read

(80) Standard English:
 a. If John had done that, ...
 b. Had John done that, ...

c. *If had John done that,...

(81) Correlation: If V can move to T, then it can continue to C (as in French); but if V can't move to T, it can't jump to C crossing T (as in English):

(82) He has not bought the book
Has he bought the book?
* He buys not the book
* Buys he the book?

(83) Il n'a pas acheté le livre
A-t-il acheté le livre?
Il n'achète pas le livre
Achète-t-il le livre?

In Early Modern English, when V to T was possible, V to (T to) C was possible too:

(84) How cam'st thou hither?
(How did you come here?)

12. Verb Second

(85) dass [Hans gestern [ein Buch gekauft] hat]
,that Hans yesterday a book bought has'

(86) [XP C [..... T]]

(87)a Hans hat [___ gestern ein Buch gekauft ___]
,Hans has yesterday a book bought'

b Gestern hat [Hans ___ ein Buch gekauft ___]
,Yesterday has Hans a book bought'

c Ein Buch hat [Hans gestern ___ gekauft ___]
,A book has Hans yesterday bought'

13. The Determiner Phrase.

(88) What is the structure of [D N]?

a. [[D] N]

b. [D [N]]

The correct structure is b: many languages manifest N to D movement, akin to V to T movement, which is expected under b but not under a.

(89)a Rumanian :
un portret
'A portrait'

b acest portret
'This portrait'

c portret-ul
'portrait+the'

- (90) Gianni ha letto [due [libri di linguistica]]
Gianni ne ha letti [due [____]]

Here *ne* pronominalizes the whole NP within the DP:

- (91) Libri di linguistica, Gianni ne ha letti [due [____]]

References.

- Adger, D. (2003) *Core Syntax – A Minimalist Approach*, Oxford University Press, Oxford-New York.
- Baker, M. (2001) *The Atoms of Language*, Basic Books, New York.
- Belletti, A. (1990) *Generalized Verb Movement*, Rosenberg & Sellier, Turin.
- Belletti, A, ed. (2004) *Structures and Beyond – The Cartography of Syntactic Structures, Vol. 3*, Oxford University Press.
- Cecchetto, C. (2002) *Introduzione alla sintassi*, LED, Milano.
- Chomsky, N. (1957) *Syntactic Structures*, Mouton, The Hague
- Chomsky, N. (1981) *Lectures on Government and Binding*, Foris Publications, Dordrecht.
- Chomsky, N. (1995) *The Minimalist Program*, MIT Press, Cambridge, Mass.
- Chomsky, N. (2003) *On Nature and Language*, Cambridge University Press, Cambridge.
- Cinque, G. (1999) *Adverbs and Functional Heads*, Oxford University Press, New York.
- Cinque, G. (2005) "Deriving Greenberg's Universal 20 and Its Exceptions", *Linguistic Inquiry*, 36, 315-332.
- Crain, S. (1992) "Language Acquisition in the Absence of Experience", *Behavioral and Brain Sciences*, 14: 597-611.
- Descartes, R. (1637/1951) *Discours de la méthode*, Union générale d'éditions, Paris.
- Dryer (1992) « The Greenbergian Word Order Correlation », *Language* 68, 81-138.
- Greenberg, J. (1963) "Some Universals of Grammar with Particular Reference to the Order of Meaningful Elements", in J. Greenberg, ed., *Universals of Language*, MIT Press, Cambridge, Mass, 73-113.
- Haegeman, L. (1994) *An Introduction to Government-Binding Theory*, 2nd edition, Blackwell, Oxford.
- Hauser, M., N. Chomsky, T. Fitch (2002) "The Faculty of Language: What is it, Who has it, and how did it evolve?" *Science* 298, 2002.
- Kayne, R. (1994) *The Antisymmetry of Syntax*, MIT Press, Cambridge, Mass.
- Lenneberg, E. H. (1967) *Biological Foundations of Language*, John Wiley & sons, New York.
- Mehler, J. & E. Dupoux (1990) *Naître humain*, Editions Odile Jacob, Paris.
- Pinker, S. (1994) *The Language Instinct*, Morrow, New York.
- Pollock, J.-Y. (1989) "Verb Movement, UG and the Structure of IP", *Linguistic Inquiry*, 20: 365-424.

- Radford, A. (2004) *Minimalist Syntax*, Cambridge University Press, Cambridge.
- Rizzi, L. (1982) *Issues in Italian Syntax*, Foris Publications, Dordrecht.
- Rizzi, L. (1990) *Relativized Minimality*, MIT Press, Cambridge, Mass.
- Rizzi, L. (1997) “The Fine Structure of the Left Periphery”, in L. Haegeman, ed. *Elements of Grammar*, Kluwer, Dordrecht, 281-338.
- Rizzi, L. (2000) *Comparative Syntax and Language Acquisition*, Routledge, London.
- Rizzi, L. (2012) “The Functional Structure of the Sentence, and Cartography”, in M. Den Dikken, ed., *Cambridge Handbook of Generative Syntax*, Cambridge University Press, Cambridge, UK, 2012.
- Rizzi, L., ed. (2013) *Lingua 130 – Special issue « Syntax and Cognition : Core Ideas and Results in Syntax »*.