Networks in Linguistic Variation[*]

## From the Chomsky Hierarchy to Parameter Schemata

Ángel J. GALLEGO

angel.gallego@uab.cat

Universitat Autònoma de Barcelona

1. Introducing the research question: Types of linguistic networks (hierarchies, clusters).

Language variation has been the the focus of much research and discussion for the last decades, the pendulum swinging from structuralist-rooted perspectives arguing that "(l)anguages can differ from each other without limit and in unpredictable ways" (Joos 1957:96) to formal models within the Chomskyan paradigm, based on the much less patent argument that syntactic structure must be, in some core sense, universal. This quasi paradox has not been resolved yet, and even staunch generativists will grant that they cannot understand why languages differ or, more basically, along which dimensions they may. Within what Baker (2010) calls "formal generative typology", variation nuances have been dealt with invoking the notion of *parameter*, originally conceived as the degree of variation of a universal principle, which Chomsky's (1981) Principle & Parameter Theory (PPT) invited us to conceive of as a two-position switch. From this stremlined perspective, children would simply have to determine the position of the switch ("fix the parameter", if you prefer) after exposure to the data until a steady state of a particular grammar is reached. Accordingly, languages are seen as *clusters* (*networks*) of properties to be determined through observable input. And, in fact, linguists have put forward different networks to capture the relevant groupings: see (1) (cf. Baker 2001, Biberauer & Roberts 2016, and references therein).

Graphs like the ones in (1) have also been used to characterize phenomena from almost any linguistic level: phonology (distinctive features), semantics (animacy / definiteness hierarchies), morphology (pronoun geometries) and syntax (see below). And the notion of network has been applied to other research areas (e.g.,
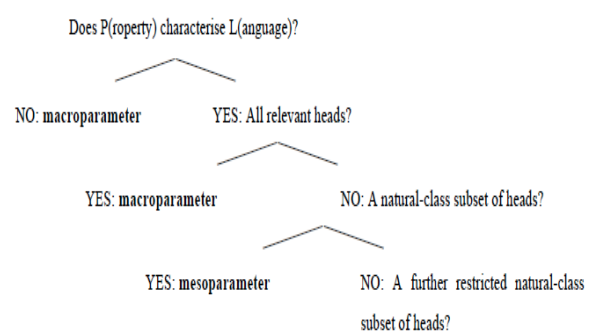
(1)

[from Biberauer & Roberts 2016:2]

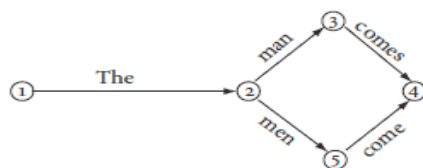economics, social networks, biology or particle physics).

Along with hierarchichal groupings like (1), linguistic theory is also well-known for providing the so-called *Chomsky Hierarchy* (CH; cf. Chomsky 1956), a theory of grammar types according to their complexity: finite-state grammars, context-free grammars and context-sensitive grammars. Given an alphabet (a finite repository of atomic units AUs), the type of objects that these grammars can generate are: lists of AUs (as in (2a)), combinations of AUs yielding abstract non-terminal units ("phrases", ❶-❷-❸ in (2b)), and the *re*-combination of phrases that are pronounced in a position where they are not interpreted in ("chains", the discontinuous ❶-❶ object in (2c)):

(2)   a. [① One], [② two], [③ three], etc.

       b. [❸ I [❷ sent [❶ e-mails]]]

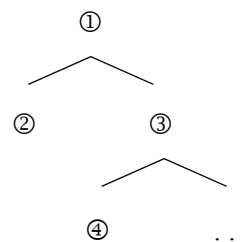   c. [❶ E-mails], [❸ I [❷ sent ❶]]]

          ↑_____|

Interestingly, the relations in (1) can be readily characterized by finite-state rules, which can yield the strings in (3). This is shown in (4), which should in turn be compared with (5) (a network that replicates (1)):

(3)   a. The old man comes

       b. The old men come

          [from Lasnik 2011:354]

(4)   Finite-state (regular) grammar rules   (5)   Standard Parameter Network

Although the logic above is clear, so is the problem: natural languages have been shown to display a mildly context-sensitive nature (cf. Chomsky 1956, Joshi 1985, Uriagereka 2008). Therefore, if (1) is supposed to be a theory of natural language variation, we should wonder why the levels of the CH do not manifest themselves there too.

2. Where the idea leads

Of course, it could be the case that our theory of language variation correlates *only partially* with the CH levels. Be that as it may, and taking the CH to be a theory with mathematical validity, here we'd

like to explore the possibility that variation and language contact facts within Iberian varieties provide evidence that a secanrio more complex than the one in (1) is actually found. That would not only be consistent with otherwise CH-compatible natural language evidence, but also with the empirically sustained observation that parametric networks display the organization in (6) (cf. Longobardi 2012), a non finite-state compatible network. Although we do not expect the kind of complexity in (6) (which goes beyond mildly context-sensitive metrics), we show evidence suggesting networks that do outrank the power of finite-state grammars.

(6)
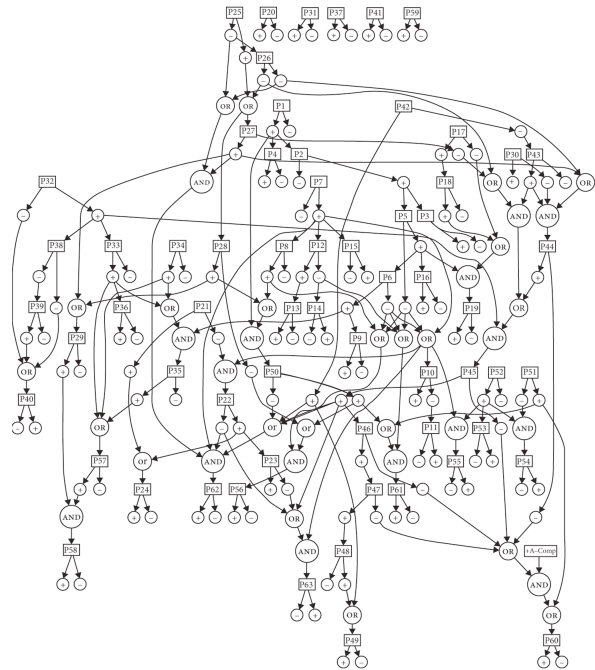
[from Longobardi 2012]

Assuming this much, this paper aims to answer to the following research questions:

(7)   a. Do the different levels of the CH manifest themselves in parameter networks?

b. If the answer to (7a) is positive, do the CH levels correlate with different types of variation?

c. What are the boundary conditions to determine the network types that account for variation?

Question (7a) speaks for itself: It is entirely an empirical issue (with solid descriptive foundations). Question (7a) has, nonetheless, key theoretical consequences, for it leads to the plausible expectation, expressed in (7b), that the different types of grammars (and networks) correspond to progressively deeper levels of linguistic variation (cf. Uriagereka 2007). Consider this possibility in (8)-(9)-(10), which shows different types of variation, implying peripheral, sub-case, and core parameters: (8) shows analogy-based (plausibly triggered by a prestigious pattern) effects like subject-verb adjacency; (9) reveals second-order evidence from Spanish clitic doubling (an option absent in many other Romance languages); and (10) shows morphological properties that allow to express arguments as agreeing elements (clitics), with consequences for additional word order facts (cf. Baker 1996).

(8)   a. I wonder [$_{CP}$ what **you did** t$_{what}$]

        b. I wonder [$_{CP}$ what **did you** t$_{what}$]

(9)   a. Alguien **(la)**   está viendo **a María**   (River Plate Spanish)

           someone cltic.her is   seeing the María

           'Someone is seeing María'

        b. Algú   **(*l')**   està veient **la Maria**   (Catalan)

           someone   cltic.her is   seeing the Maria

           'Someone is seeing María'

(10)   a. ∅ Maite ∅ nuen   (Basque)

            love   1sg.aux.3sg

          'I love her'

        b. *(I) love *(her)

Finally, (7c) is a more general question about the types of networks that natural languages can manifest (cf. Solé et al. 2010). It is here where, apart from linguistic evidence, interdisciplinary research is needed in order to see if there are non-language specific factors (cf. Chomsky 2005) that regulate linguistic networks.

There are further questions that the answers to those in (7) may help understand better:

(11)   a. If parametric networks are category-centered (verbs, determiners, tense morphemes, etc.), do we expect for "languages" to be defined as a collection of different points of different networks?

        b. Given the rigidity of parameter networks, how can they capture unexpected inter-linguistic correlations between languages that have no phylogenetic relation (e.g., Spanish and Serbo-Croatian).


3. Theoretical Consequences

The working hypothesis of this paper can be expressed as (12):

(12)   *The Chomsky Hierarchy Paramater Network Hypothesis (CHPaNet)*

        Parameter networks of linguistic variation reflect the levels of the Chomsky Hierarchy

The CHParNet predicts parametric scenarios that exceed (1), repeated below as (13). More precisely, (12) predicts situations that can be captured by both determinstic and non-deterministic algorithms:
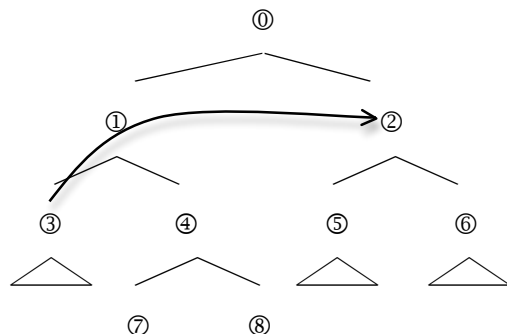
(13)
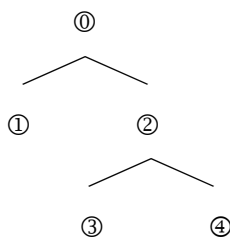
(14)
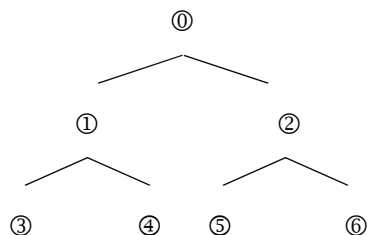
(15)

As already noted, (13) is compatible with a finite-state grammar going from one state to another one. (14) and (15) are not, as they display networks that are both left- and righ-recursive, creating (potentially) multiple learning paths—to be dertermined by linguistic input. More interestingly, (14) and (15) require non-derterministic algorithms, which have been regarded as problematic on acquisition grounds (cf. Biberauer & Roberts 2015).

Technically, the patterns that emerge in (14) and (15) are precisely those involving "multiple cascades" (or independent derivational workspaces). Given a recursive structure-building operation that takes two objects and generates a new one (call it *Merge*; cf. Chomsky 1995 and ss.), only (13) allows a derivation that does not require abandoning a single derivational workspace.

To be sure, more complex situations (like (14) and (15)) raise non-trivial morphological (agreement), semantic (binder-bindee / argument-predicate asymmetries), syntactic (transformations), and even phonetic (linear order, OCP effects) nuances, but... *they exist*. In linguistics, these domains are well-known, as they correspond to "complex left branch" (specifier) situations, which have been the locus of empirical conundrums since the early eighties (cf. Chomsky 1981, Huang 1982, Rizzi 1990, Kayne 1994, Uriagereka 1999): EPP and ECP effects, the Subject and Adjunct Conditions, LCA, multidominance, incorporation, Case assignment, agreement, binding, etc. Consider cases of

EPP, ECP and Subject Condition effects, all of which instantiate a complex left-branch situation (in (17) and (18) *t* stands for a trace left by *who*):

(16)   a. *Spoke someone

          b. Someone spoke

(17)   a. *Who did you say [$_{CP}$ that t_{who} spoke]?

          b. Who did you say [$_{CP}$ t_{who} spoke]?

(18)   a. *Who did [$_{NP}$ a critic of t_{who}] see you?

          b. *Who did you see [$_{NP}$ a critic of t_{who}]?

These matters have been at the forefront of linguistic theory for decades, but no satisfactory (let alone agreed upon, and even less so unitary) answer has been reached. The CHParNet is the hypothesis that, just like (14) and (15) depict objects that are present in natural languages (and therefore there must be a theory accounting for this; cf. Chomsky 1956, Chomsky 1995), *some aspects of this very theory plausibly carries over to our theory of language variation*. The overall discussion has an important, and thought-provoking, theoretical burden, but given the empirical robust evidence for the CH (namely (13), (14) and (15)), its absence in the context of language variation would be unexpected. Along with these linguistic concerns, the CHParNet has clear implications for research questions involving transversal notions like computational complexity, locality, globality, efficiency, etc.

4. Empirical Issues

A remarkable trait of PPT-rooted parameter networks is that they involve what we may call *Bifurcated Choice States* (BCS; states with two-option choices), even though the Network Theory literature has shown that the patterns can be more complex—"free-scale" networks, vis-à-vis "random" ones (see (19)).

      BCS lead to scenarios whereby a given choice at a state S does not allow "backtracking", thus barring dependencies with the offsprings of an already passed state in the same cascade or in a different one: *Multifurcated Choice States* (MCS). However, that is precisely what we expect if the CHParNet holds.

(19) a.

RANDOM NETWORK

b.

FREE-SCALE NETWORK
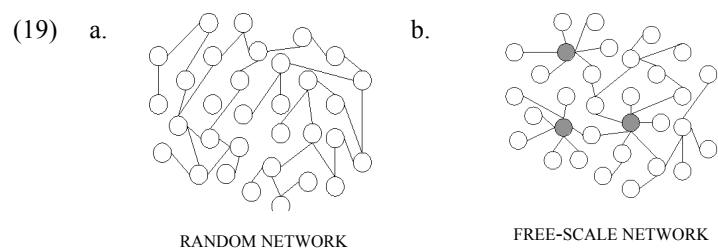
To empirically substantiate the CHParNet, let us consider linguistic evidence displayed by Iberorromance varieties that illustrate, we claim, MCS. To be specific, let us discuss a series of Verb-Object morpho-syntactic phenomena that are restricted to the VP:

(20)   Verb-Object morpho-syntactic phenomena

      a. VOS order   e. Case-displacement

      b. VSO order   f. Participial agreement

      c. Differential Object Marking   g. Auxiliary selection

      d. Oblique clitics   h. Possessive *have*

The range of variation of the phenomena in (20) show that they are connected in interesting ways and give rise to a continuum that goes from strongly configurational Romance languages of the French type to partially configurational Romance languages of the Spanish type, with languages such as Catalan and Italian being along the way. There are various aspects that must be clarified at this point: first, why we focus on Iberorromance languages (taking Spanish to play a key role); second, how the facts in (20) are connected; and third, in what sense these pieces of evidence support the CHParNet.

The phenomena listed in (20) reveal deeper grammatical bonds beyond their obvious local character (they belong to the VP domain; cf. Kayne 1989, Picallo 1998, Fernández-Ordóñez 1999, Costa 2000, Ordóñez 2000, Torrego 2002, Belletti 2005, D'Alessandro 2007, Lopez 2012, Gallego 2013, Zubizarreta 1998). Before seeing that, consider a restricted sample of some of the facts in (20):

(21)   Iberorromance languages use different strategies to derive VOS:
      Object shift or VP dislocation

      a. Cogió   **cada coche**$_i$ su$_i$ dueño (Spanish)   propietari   (Catalan)

         took-3.sg each car   its owner

         'Its owner took each car'

      b. *?Va   agafar   **cada cotxe**$_i$ el seu$_i$ (Catalan)

          aux-3.sg take-inf each car   the its   owner

         'Its owner took each car'

(22)   Spanish (E.Portuguese, Galician, and Basque) license VSO sentences, but Catalan does not
      a. Hoy   compró **Juan** el diario   (Spanish)

         today buy-3.sg Juan the newspaper

         'Juan buys the newspaper everyday'

      b. *Fullejava   **en Joan** el diari   (Catalan)

         browsed-3.sg the Joan   the newspaper

         'Joan was browsing the newspaper'

(23)   Spanish displays DOM, while Basque does only in dialects in contact with Spanish

    a. Hillary saludó   **a Donald**   (Spanish)

       Hillary greeted.3.sg   dat Donald

       'Hillary greeted Donald'

    b. Peruk   **Joneri**   ikusi   dotzo   (Lekeitio Basque)

       Peru.erg Jon.dat   see   aux.dtrn

       'Peru saw Jon'

If we put all these pieces together, we get (24) (% signals that the phenomenon is subject to dialectal restrictions).

In certain respects, Spanish aligns more with Portuguese, Galician and Basque than it does with Catalan (which is in turn closer, empirically and geographically, to, e.g., French and Italian). This indicates that spatial factors (language contact) play a role in determining variation. At the same time, and although (24) does not reflect it, many of the phenomena present-day Spanish does not license (oblique clitics, participial agreement, auxiliary selection, etc.) *did show up* in earlier stages. In (25) we see this in the case of auxiliary selection and participial agreement, which were possible in Old Spanish.

(24)

IBERIAN LANGUAGES

	E.POR.	GAL.	E.SPA.	CAT.	BAS.
1. VOS (via OS)	☑	☑	☑	☒	☑
2. VSO	☑	☑	☑	☒	☑
3. DOM	☒	☒	☑	☑ (%)	☑ (%)
4. Oblique clitics	☒	☒	☒	☑	☒
5. Case displacement	☒	☒	☑	☒	☑ (%)
6. PPart. agreement	☒	☒	☒	☑ (%)	☒
7. Auxiliary selection	☒	☒	☒	☒	☑
8. Possessive HAVE	☒	☒	☒	☒	☑ (%)

(25)   a. Los dineros   los   avia   **mandados**   (Old Spanish)

       the money-m.pl they-m.pl had-3.sg sent-part.m.pl

       'And the (pieces of) money he had send'

           [XIIIth century, from CORDE database]

    b. Quando sopo   que su hermana **era** ida   (Old Spanish)

       when   knew-3.sg that her sister   was left

       'When she knew that her sister had left'

           [from Aranovich 2003: 4]

Diachronic evidence is not the main target of this proposal, but the fact that the very phenomena we want to study are now absent in Spanish (though not in Catalan, French and Italian), but were not

centuries ago indicates that deeper correlations are at hand—deeper and, again, *more complex*, in accord with the CHParNet.
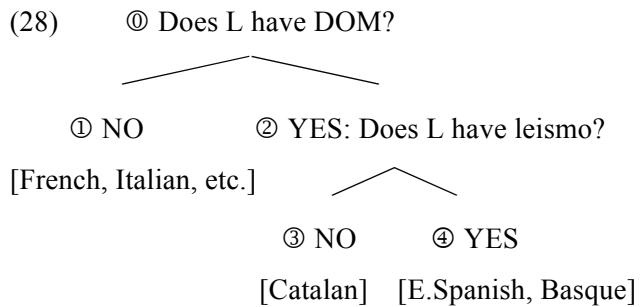
Unlike English, all Iberian languages can drop their arguments (they are "pro-drop"). Nonetheless, and although the observations in the previous section indicate that language contact is an important factor driving variation, there are certain pieces of evidence that suggest that the types of connections needed to capture the facts favor the CHParNet. Let us concentrate on two phenomena that go hand-in-hand: differential object marking (DOM) and a particular instance of Case displacement, so-called "leismo".

As (23) showed, DOM (the insertion of a dative Case marker before the direct object: *a*) is present in Spanish and in some varieties of Basque (precisely those in contact with Spanish). The case of Catalan is murkier, for although normative Catalan is said to rule out DOM, colloquial varieties (especially those where the contact with Spanish is stronger, like in the Barcelona province) do deploy DOM. With this in mind, consider leismo, which has been analyzed as a process of dativization (alternatively, accusative Case displacement) of the direct object by Colomina et al. (2016), who explicitly connect the dative Case marker of DOM to leismo (e.g. *Admiro **a Barack*** ('I admire to Barack') → ***Le** admiro* ('I admire him$_{DAT}$')). Basque nicely fits with this approach, since the agreement markers in the verb are dative (see (26) below). However, no variety of Catalan resorts to dative clitics, which are equivalent to the dative agreeing morphemes of Basque (see (27), where evidence comes from Barcelonese Catalan):

(26)   Nik   **suri**   ikusi dotzut   (Lekeitio Basque)

I.erg   you.dat seen   3sgA.**2sgD**.1sgE

'I've seen **to you**' (Eng. 'I've seen you')

(27)   Estimo   **al**   **Pere** → ***Li***   estimo   (Barcelonese Catalan)

love-1.sg to-the Pere   him$_{DAT}$ love-1.sg

'I love **to Pere**' (Eng. 'I love Pere')

What is intriguing here, to repeat, is that although some varieties of Catalan resort to *bona fide* DOM, leismo is totally out. Why, if one of them clearly seems to biggy-back on the other? This suggests that even quite closely-related languages in contact situations are differentially permeable to borrowing certain grammatical properties and not others. Plausibly, the un-borrowed parameter (leismo) is more complex than the borrowed one (DOM). In terms of standar parameter approaches, leismo must be more buried within the network than DOM. Suppose that, somewhat impressionistically, we formalize this as (28). The problems of such parameter network should be clear:

On the one hand, the network links two phylogenetically un-related languages (Spanish and Basque) in a very intimate fashion, leaving Catalan out. On the other, Catalan is far from other varieties (French and Italian) with which it aligns for the purposes of a bunch of VP phenomena: oblique clitics, participial agreement, and lack of VSO.

(28) ⓪ Does L have DOM?

① NO   ② YES: Does L have leismo?
[French, Italian, etc.]

③ NO   ④ YES
[Catalan]   [E.Spanish, Basque]

In a nutshell, this restricted sample is enough to reveal the limitations (and *prima facie* contraditions) of classical parameter networks. Granted, the problem may well lie in the fact that the phenomena in (20) are regarded as primitives, but even with that proviso in mind, the outcome in (28) imposes an unwanted rigidity on the connections. The appeal of the CHParNet is that those connections are still possible in a more dynamic fashion.

References

Baker, M.C. 2001. *The Atoms of Language*. Basic Books: New York.

Baker, M.C. 2010. "Formal Generative Typology". In B. Heine and H. Narrog (eds.), *The Oxford Handbook of Linguistic Analysis*, 285-312. Oxford: Oxford University Press.

Biberauer, T. and I. Roberts. 2015. "Rethinking Formal Hierarchies: A Proposed Unification". *Cambridge Occasional Papers in Linguistics* 7: 1-31.

Chomsky, N. 1956. "Three models for the description of language". *IRE Transactions on Information Theory* 2: 113-124.

Chomsky, N. 1981. *Lectures on Government and Binding*. Dordrecht: Foris Publications.

Joos, M. 1957. *Readings in Linguistics: The Development of Descriptive Linguistics in America since 1925*. Washington DC: American Council of Learned Societies.

Kayne, R. 1994. *The Antisymmetry of Syntax*. Cambridge, MA: MIT Press.

Lasnik, H. 2011. "What kind of computing device is the human language faculty?". In A. M. di Sciullo and C. Boeckx (eds.), *The Biolinguistic Enterprise: New Perspectives on the Evolution and Nature of the Human Language Faculty*, 354-365. Oxford: Oxford University Press.

Longobardi, G., 2012, "Convergence in parametric phylogenies: homoplasy or principled explanation". In Ch. Galves, S. Cyrino, R. Lopez, and J. Avelar (eds), *Parameter Theory and Linguistic Change*, 304-319. Oxford: Oxford University Press.

Uriagereka, J. 1999. "Multiple spell-out". In N. Hornstein and S. Epstein (eds.), *Working Minimalism*, 251-282. Cambridge, MA: MIT Press.

Uriagereka, J. 2008. *Syntactic Anchors. On Semantic Structuring*. Cambridge: Cambridge University Press.