On the directionality of Non-Local Dependencies

Cristiano Chesi (cri@mit.edu) - University of Siena; MIT

this work is part of my Ph.D. thesis, available soon at:
http://www.ciscl.unisi.it/persone/chesi.htm

Abstract

Minimalist Inquiry (Chomsky 1995-2001) proposes a model of linguistic competence able to build phrase structures from **bottom-to-top** in a derivational way (namely starting from the inner verbal shell, up to the matrix one, adding piecemeal arguments and functional elements).

Including structure building operations (such as **merge**, **move** and the idea of **derivation by phase**) as part of the grammar has been a necessary move toward a cognitively/computationally/formally more plausible model. On the other hand, this forces us to consider the **competence** from a **procedural perspective** rather than simply from a static **knowledge-representation** point of view. Even though much emphasis is put on the fact that the proposed model is just abstract, then that no implications at all should be drawn from any performance perspective, it is worth wondering to which extent this model of competence can be accessed by performance tasks (rather than simply refer to underspecified "interface conditions").

One of the most puzzling issues from this perspective lies behind the directionality of the phrase structure building process: many cues point out that both in comprehension (**garden path** sentences) and in production (**false starts**) the process is **incremental** and it happens **from left to right** (or **top-down**) but very few attempts of building minimalist models that keep this factor into account have been explored. Notably, Phillips (1996) work goes in this direction, showing how incrementally building structure from left to right can be justified both empirically (**Contradictory Constituency**) and cognitively (**branch right** predictions). Phillips' work, however, mostly deal with local relations (it significantly redefine the ideas of **merge** and constituency formation).

This poster aims to show that this intuition is productively viable even from a **Non-Local Dependencies** perspective, namely the notion of **movement** (but probably even **pronominal binding** and **control** could be captured with the very same idea, cf. Chesi 04) could be radically redefined within a **left-to-right** (**top-down**) framework, without loosing empirical power, gaining, indeed, computational efficiency (the proposed solution is **tractable**) and cognitive plausibility (**no look-head** or **abstract un-interpretable features** are required). In order to do that we should define:

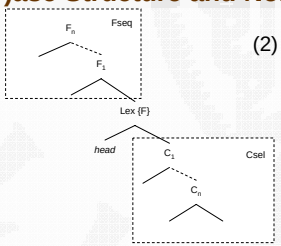
1 - the formal context that allows for the definition of Non-Local Dependencies (**Grammar Specification** and **Locality Conditions**);

2 - the nature of the process during which these relations are established (**Derivation by Phase**);

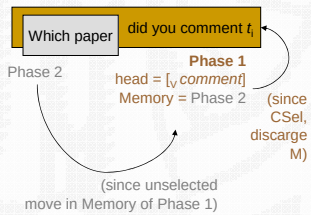
3 - what triggers these Dependencies (**Selection Requirements**).

P(r)ase Structure and Non-Loc. Dependencies

(1)



(2)



1 - Grammar Specification and Locality

Minimalist Grammars (Chomsky 95-01) formalizations (Stabler 97) can be defined as **4-tuples** {**V**, **Cat**, **Lex**, **F**} such that:

V is a finite set of non-syntactic features (**phonetic** and **semantic** features),

Cat is a finite set of syntactic features composed by, among others, **selectional** features (expressing the complement(s) requirement(s) of the element bearing them), **licensors** features (following Chesi's 04 formalization, an ordered list of functional specifications inspired to cartographical inquiries, Cinque ed. 2002 and related work) and **interveners** (a set of classes conformed to relativized minimality intervention effects, cf. Rizzi 90, 04),

Lex is a finite set of expressions built from **V** and **Cat** (the lexicon)

F is a set of partial functions from tuples of expressions to expressions (minimally speaking **merge** and **move**, plus, following Chesi 04, a **phase projection** device, see box below).

Assuming a simplified phrase structure (**doing away with specifiers**, Starke 01) and accepting a **linearization** idea (Kayne 94), let us call **F(functional)Seq(uence)** the ordered list of **licensors** on the left of any lexical head and **C(complement) Sel(ection)** the ordered list of obligatorily selected complements on the right of any given lexical head as shown in (1).

2 - Phase (Projection)

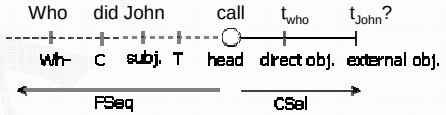
Assume a **Phase** to be the minimal part of a top-down computational process in which all **functional** and **selectional** specifications associated to a given lexical head (**N** or **V**) are satisfied (then **CP** and **DP** are phases, but crucially not **vP**).

Formally speaking, a phase is a **Top-Down expectation**, that is, the projection of a set of **merge** relations among underspecified (at least in terms of π features) lexical items. This set is **finite** since it contains exactly **one** projecting lexical element (by definition of phase), **+** (at worst) **three** ph(r)ases (each lexical head **selects** at most 3 arguments, Pesetsky 1982) **+ K** functional specifications (Cinque 02, even though empirical evidence points out that no more than 4, 5 functional specifications can be processed at the same time, cf. Chesi 04). A Top-Down expectation can be (partially) refined during the derivation, according to the new items merged.

3 - Non-Local Dependencies as Unsatisfied Selections

Assume the sentence structure to be described by a finite set of **Lexical Items Identifiers** (**inclusiveness condition**, Chomsky 95) among which a finite set of **immediate dominance** (that is, holding only between adjacent items: this way dominance corresponds to **merge**) and **immediate precedence** relations are defined.

Since any lexical element plus the grammar specification provide enough information about the geometry of any item allowed around the lexical entry that originates the phase, any "mis-placed" element should be considered a case of Non-Loc. Dependency. This is because of the coexistence, on the same lexical element of features that are licensed in non-contiguous places, e.g. *patient thematic role* (to be considered, for instance, as the first **CSel** requirement of a verbal head) and *wh-/question* feature (that is, a **FSeq** specification of the verbal phase):



Capturing Non-local dependencies

In order to establish a Non-Local Dependency from a left to right (top-down) perspective, we need to:

1. find out that an element is in a non-selected position;
2. "bring with us", during the processing, an expectation about the selectional requirement to be satisfied in order to legitimate (by re-merge) the actual element position;
3. consider any possible intervening element that could block this non-local relation.

To "implement" this intuition we can postulate a **memory buffer** (somehow reminiscent of the **hold** register of ATNs, Woods 70, or Fong's 04 minimalist implementation of this device) in order to store the moved element as long as the relevant selecting head is found. So as to capture the correct locality constraints we should postulate that this memory buffer has a **slot-structure** (each slot corresponding to the interveners class of the present features according to the **interveners** class specified in the grammar) and that the elements are stored and retrieved from these slots in a **Last In First Out** way. Because of complexity issues (Chesi 04) this mechanism would be unusable unless bounded in their domain of application. This domain corresponds to the revised notion of **phase** presented here.

A **success condition** requires a successful **re-merge** (that is a full **unification**, in Shieber's 86 terms, between the projected position and the element in memory) that discharge the buffer; moreover, at the end of each phase, the local buffer should be empty, or else its content will be "inherited" by the memory buffer of the (last) selected phase. A successful inheritance produces an (intermediate) trace on the edge of the phase so as to mark the presence of the element within the new chunk of processing (coherently with the **phase balance** idea, Felser 01); see (2) for a simple example of derivation (a verbal phase, Phase 1, is assumed to be the default Top-Down expectation when the derivation starts).

This model is empirically adequate in the following contexts:

- **Strong Island conditions / Connectedness effects / Parasitic Gaps** (both moved subjects and adjuncts, since un-selected, cannot discharge the memory buffer of the matrix clause, this cause the ungrammaticality of sentences that do not present any "gap" in the matrix verbal phase);
- **A(rgumental) movement / control** (passives, raising and control verbs do not pose any problem if we assume that the verbal features that identify these classes force the inspection of the memory buffer before parsing/producing new items);
- **Cross-Linguistic Variation** (phenomena such as Japanese SOV, scrambling and Warlpiri apparent "free" word order, indeed, could use case markers to trigger the correct Top-Down expectations about the ordered verbal CSel requirements).

References

Chesi C. (2004) *Phases and Cartography in linguistic computation: toward a cognitively motivated computational model of the human linguistic competence*. Ph.D. Thesis. University of Siena
Chomsky N. (1995) *The Minimalist Program*, MIT Press, Cambridge, MA.
Chomsky N. (2001) *Beyond explanatory adequacy*. MIT Occasional Papers in Linguistics, no. 20. Cambridge, MA
Cinque G. ed. (2002) *The Structure of DP and IP. The Cartography of Syntactic Structures*, vol. 1, OUP.
Felser C. (2001) *Wh-Copying, Phases and Successive Cyclicity*. Essex Research Reports in Linguistics, 37
Fong S. (2004) *Computation with Probes and Goals: A Parsing Perspective*. In Di Sciullo, A. M. and R. Delmonte Ed. UG and External Systems. Amsterdam: John Benjamins. 2004 (In press)
Kayne R. (1994) *The antisimmetry of syntax*. MIT Press, Cambridge, MA.
Phillips (1996) *Order and Structure*. Ph.D. Thesis. MIT
Rizzi L. (1990) *Relativized Minimality*, MIT Press, Cambridge, Mass.
Rizzi L. (2004) On the form of chains. Ms. University of Siena.
Shieber S.M. (1986). *An Introduction to Unification-Based Approaches to Grammar*. Stanford, CA: CSLI.
Stabler E. (1997) *Derivational minimalism*. in Retoré, ed. Logical Aspects of Computational Linguistics. Springer
Starke M. (2001) *Move Dissolves into Merge*. Ph.D. Thesis, University of Geneva.
Woods W. A. (1970) *Transition Network Grammars for Natural Language Analysis*. CACM 13,10: 591-605