Developments of Liner Logic for the Study of Language and Cognition

Roma 09 - 07 - 2010

Claudia Casadio

University "G. D'Annunzio" - Chieti

casadio@unich.it

Contents

1 Summary 4
2 The Language Faculty 5
3 Cognitive Categories 7
3.1 Syntactical vs. sentential 7
3.2 Semantical vs. lexical 7
3.3 Semantical vs. propositional 8
3.4 Dircourse and pragmatics 8
3.5 Cognitive operations 8
3.6 Examples 9
4 Type Logical Grammars 10
4.1 Syntactic Calculus 10
4.2 Non-commutative Linear Logic 13
4.3 Compact Bilinear Logic 15
4.4 Adjoints 16
4.5 Pregroups 17
4.6 Definition of pregroup 18
5 Computations and Short Term Memory 20
6 Language 21

6.1 Lexicon 21
6.2 Features 22
6.3 Templates 23
7 Binding 24
7.1 Wh-pronouns 24

1. Summary

Human language and cognition are related in several ways, as shown by the large literature in the field (e.g. Chomsky 2005, Jackendoff 2002, Jackendoff & Pinker 2005). These two faculties, as they are frequently referred to, are here considered in the light of the developments of Linear Logic, paying particular attention to cognitive categories and the role played by semantic vs. short term memory. A brief insight is also proposed into human reasoning.

2. The Language Faculty

In our view, syntax functions in the overall system of language as a regulator: it helps determine how the meanings of words are combined into the meanings of phrases and sentences. Syntax employs at least four combinatorial devices.

S. Pinker, R. Jackendoff, The faculty of language: what's special about it?, 2005

- **The first collects words hierarchically into syntactic phrases, corresponding (in prototypical cases) to constituents of meaning.**
- **The second orders words or phrases within a phrase, for example, by specifying that the verb of a sentence falls in a certain position such as second, or that the phrase serving as the topic comes first. Most languages of the world are not as strict about word order as English, and often the operative principles of phrase order concern topic and focus, a fairly marginal issue in English grammar.**

- **A third major syntactic device is agreement, whereby verbs or adjectives are marked with inflections that correspond to the number, person, grammatical gender, or other classificatory features of syntactically related nouns.**
- **The fourth is casemarking, whereby noun phrases are marked with inflections (nominative, accusative, and so on) depending on the grammatical role of the phrase with respect to a verb, preposition, or another noun.**

Different languages rely on these mechanisms to different extents to convey who did what to whom, what is where, and other semantic relations. English relies heavily on order and constituency, but has vestigial agreement and no case except on pronouns. The Australian language Warlpiri has virtually free word order and an exuberant system of case and agreement; Russian and Classical Latin are not far behind. Many languages use the systems redundantly, for instance German, with its rich gender and case systems, moderate use of agreement, and fairly strong constraints on phrase order.

3. Cognitive Categories

3.1. Syntactical vs. sentential

- **subject : noun argument in nominative case – NOM**
- **(direct) object : verb argument in accusative case – ACC**
- **(indirect) object : verb argument in oblique case – OBL**
- **dative : verb argument in dative case – DAT**
- **locative : verb argument in oblique case – OBL, LOC**

3.2. Semantical vs. lexical

- **gender : Feminine, Masculin, Neuter, ..., living thing, substance, ...**
- **number : Singular, Plural, Dual, ..., Collective, Discrete, ...**
- **person : 1st, 2nd, 3rd, 4th, 5th, 6th, ...**

3.3. Semantical vs. propositional

- **agent : active dimension of an action/event/proposition – NOM**
- **patient : recipient dimension of an action/event/proposition – ACC, OBL**
- **referee : goal dimension of an action/event/proposition – DAT, OBL**

3.4. Dircourse and pragmatics

- **topic : given, known argument – OLD**
- **focus : new, not (or not totally) known argument – NEW**

3.5. Cognitive operations

- **order : precedence, linearity**
- **dominance : before, after, under, within, ...**
- **binding : scope, coordination, anaphora, ...**

3.6. Examples

- **case assignment : Peter believes her to be beautiful**
- **agent vs. patient : Mary is adored by Peter**
- **given vs. new :**
  Peter knows her very well;
  Pietro la conosce bene;
  Pietro conosce lei meglio di lui .
- **dominance : His mother asked Peter for a cup of tea.**
- **wide scope vs. narrow scope : quantifiers**

  Every man (loves some woman) :
  $\forall$(x)((man: x) $\rightarrow$ $\exists$(y)((woman: y) $\wedge$ (love x y)))

  (Every man loves) some woman :
  $\exists$(y)((woman: y) $\wedge$ $\forall$(x) ((man: x) $\rightarrow$ (love x y)))

4. Type Logical Grammars

4.1. Syntactic Calculus

The Syntactic Calculus (Lambek 1958, 1961), developed from earlier systems by Ajdukiewicz and Bar-Hillel, allows compound types to be built up, starting from a finite set of basic types, by three operations:

$$\mathbf{A} \otimes \mathbf{B} \ , \ \mathbf{A}/\mathbf{B} \text{ and } \mathbf{A}\backslash\mathbf{B} \ ,$$

subject to the following axioms and rules of inference:

$$(\mathbf{AB})\mathbf{C} \leftrightarrow \mathbf{A}(\mathbf{BC}) \ ,$$

$$\mathbf{AB} \rightarrow \mathbf{C} \ \text{ iff } \ \mathbf{A} \rightarrow \mathbf{C}/\mathbf{B} \ \text{ iff } \ \mathbf{B} \rightarrow \mathbf{A}\backslash\mathbf{C} \ ,$$

where $\rightarrow$ is a deduction symbol satisfying reflexivity and transitivity.

We can abbreviate $\mathbf{A} \otimes \mathbf{B}$ as AB.

In the Syntactic Calculus you prove theorems like:

$(A/B)\ B \rightarrow A$	**Ajdukiewicz's law ,**
$B \rightarrow (A/B)\backslash A$	**type raising ,**
$(A/B)(B/C) \rightarrow A/C$	**composition ,**
$A/B \rightarrow (A/C)/(B/C)$	**Geach's law ,**
$(C/B)/A \leftrightarrow C/(AB)$	**Curry's law ,**

as well as their mirror images such as:

$$B\ (B\backslash A) \rightarrow A \quad \text{etc.}$$

We have also

$$\mathbf{(A\backslash B)/C \leftrightarrow A\backslash(B/C)}\ ,$$

hence we often write $\mathbf{A\backslash B/C}$ **for either side,**

and the following rules of inference indicating what categorists call functoriality:

$$\frac{\mathrm{A \to B \quad C \to D}}{\mathrm{AC \to BD}} \qquad \frac{\mathrm{A \to B \quad C \to D}}{\mathrm{A/D \to B/C}}$$

and the mirror dual of the latter.

4.2. Non-commutative Linear Logic

After Girard (1987), both Abrusci (1991) and Lambek (1993) develop non-commutative systems of linear logic

non-commutative linear logic (Abrusci)

classical bilinear logic (Lambek).

The second is obtained from the Syntactic Calculus by adjoining a constant 0 satisfying

$$0/(A\backslash 0) \leftrightarrow A \leftrightarrow (0/A)\backslash 0$$

that can also be written

$$A\backslash 0 = A^r, \quad 0/A = A^\ell \ .$$

Abrusci, V. M. Phase semantics and sequent calculus for pure Noncommutative Classical Linear Propositional Logic. The Journal of Symbolic Logic, 56: 1403–1451, 1991.

Lambek, J. From categorial grammar to bilinear logic. In K. Došen and P. Schroeder-Heister (eds.), Substructural Logics, 207-237, Oxford University Press, 1993

One can show that

$$(\mathsf{B}^r\,\mathsf{A}^r)^\ell \leftrightarrow (\mathsf{B}^\ell\mathsf{A}^\ell)^r\ ,$$

for which we can also write (A+B) (equivalent to $\mathsf{A}\wp\mathsf{B}$). Here are some theorems of Classical Bilinear Logic, anticipated by Grishin (1983):

$$\begin{array}{rcl}
1^r \leftrightarrow 0 \leftrightarrow 1^\ell & & \\
\mathsf{A}{+}0 \leftrightarrow \mathsf{A} \leftrightarrow 0{+}\mathsf{A} & & \\
(\mathsf{A}{+}\mathsf{B}){+}\mathsf{C} \leftrightarrow \mathsf{A}{+}(\mathsf{B}{+}\mathsf{C}) & & \\
\mathsf{A}^\ell\mathsf{A} \to 0 & , & \mathsf{A}\mathsf{A}^r \to 0 \\
1 \to \mathsf{A}{+}\mathsf{A}^\ell & , & 1 \to \mathsf{A}^r{+}\mathsf{A} \\
\mathsf{A}/\mathsf{B} \leftrightarrow \mathsf{A}{+}\mathsf{B}^\ell & , & \mathsf{B}\backslash\mathsf{A} \leftrightarrow \mathsf{B}^r{+}\mathsf{A} \\
(\mathsf{A}{+}\mathsf{B})\,\mathsf{C} \to \mathsf{A}+\mathsf{B}\mathsf{C} & , & \mathsf{C}\,(\mathsf{B}{+}\mathsf{A}) \to \mathsf{C}\mathsf{B}+\mathsf{A}
\end{array}$$

Grishin V., On a generalization of the Ajdukiewicz-Lambek system, Nauka, Moscow (1983); in Abrusci and Casadio (eds.), Proceedings 5th Roma Workshop, Bulzoni Editore, Roma (2002).
Moortgat M., Symmetric Categorial Grammar, J. Philos Logic, 38, 681-710, 2009.

4.3. Compact Bilinear Logic

Lambek argues that there is a simplification if you assume

$$\mathbf{A+B} \leftrightarrow \mathbf{AB} \quad , \quad \mathbf{0} \leftrightarrow \mathbf{1} \ .$$

The word "compact" had been used e.g. by Barr (1979) to describe this situation in a categorical context, though then still restricted to the commutative case. Finally, it was realized that compact bilinear logic allowed a simpler description as follows:

$$\mathbf{(AB)C} \leftrightarrow \mathbf{A(BC)} \quad , \quad \mathbf{A1} \leftrightarrow \mathbf{A} \leftrightarrow \mathbf{1A}$$
$$\mathbf{A}\,\mathbf{A}^r \to \mathbf{1} \to \mathbf{A}^r\mathbf{A} \quad , \quad \mathbf{A}^\ell\,\mathbf{A} \to \mathbf{1} \to \mathbf{A}\,\mathbf{A}^\ell \ .$$

Barr, M. (1979). $*$-Autonomous Categories. Springer LNM 752, Berlin.

4.4. Adjoints

Models in which $\rightarrow$ stands for a partial order are called "pregroups";

they reduce to groups if the order is discrete, that is, is the equality relation, and to partially ordered groups in the so-called cyclic case when $A^\ell \leftrightarrow A^r$.

If however the arrows are allowed to stand for morphisms in a category, one would also demand that the above occurrences of $\leftrightarrow$ represent isomorphisms and that the composite arrows

$$A \rightarrow A\,A^\ell A \rightarrow A \quad , \quad A \rightarrow A\,A^r A \rightarrow A$$

are identity arrows, making A^ℓ the left adjoint and A^r the right adjoint of A in a 2-category.

4.5. Pregroups

Type logical grammars developed from Compact Bilinear Logic, with two distinctive features

one-dimensional representation

checking grammaticality by simple calculations

Whereas type-logical grammars treat syntactic derivations as logical proofs, usually represented by two-dimensional diagrams, I here wish to defend the view that people process linguistic information by one-dimensional calculations ... and will explore an algebraic approach based on the notion of a "pregroup", a partially ordered monoid in which each element has both a left and a right "adjoint".

... one assigns to each word one or more "syntactic types", elements of the free pregroup generated by a partially ordered set of "basic types", grammaticality of a string of words is checked by a calculation on the corresponding types. Calculations are performed according to Miller's definition of temporary storage capacity of our short-term memory.

Lambek, J., A computational algebraic approach to English grammar, Syntax, 7(2), 128-147, 2004.

4.6. Definition of pregroup

A pregroup $\{G, \cdot, 1, {}^{\ell}, {}^{r}, \to\}$ is a partially ordered monoid in which each element a has a left adjoint a^{ℓ}, and a right adjoint a^{r} such that

$$a^{\ell} a \to 1 \to a\, a^{\ell}$$
$$a\, a^{r} \to 1 \to a^{r} a$$

the dot "." stands for multiplication with unit 1, and the arrow denotes the partial order. In linguistic applications the symbol 1 is the empty string of types and multiplication is interpreted as concatenation. Adjoints are unique and you prove

$$1^{\ell} = 1 = 1^{r},$$

$$(a \cdot b)^{\ell} = b^{\ell} \cdot a^{\ell} \quad , \quad (a \cdot b)^{r} = b^{r} \cdot a^{r},$$

$$\frac{a \to b}{b^{\ell} \to a^{\ell}} \quad , \quad \frac{a \to b}{b^{r} \to a^{r}} \quad , \quad \frac{b^{\ell} \to a^{\ell}}{a^{\ell\ell} \to b^{\ell\ell}} \quad , \quad \frac{b^{r} \to a^{r}}{a^{rr} \to b^{rr}} \quad .$$

The following also hold

$$\mathbf{a}^{r\ell} = \mathbf{a} = \mathbf{a}^{\ell r} \ ,$$
$$\mathbf{a}^{\ell\ell}\ \mathbf{a}^{\ell} \to \mathbf{1} \to \mathbf{a}^{\ell}\ \mathbf{a}^{\ell\ell} \ ,$$
$$\mathbf{a}^{r}\ \mathbf{a}^{rr} \to \mathbf{1} \to \mathbf{a}^{rr}\ \mathbf{a}^{r} \ ,$$

We work with the pregroup freely generated by a partially ordered set of basic types. From each basic type a we form simple types by taking single or repeated adjoints:

$$\ldots\ a^{\ell\ell},\ a^{\ell},\ a,\ a^{r},\ a^{rr}\ldots$$

Compound types or just types are strings of simple types. We assign to each word or word form in the dictionary of the language under investigation one (or more) types. The only computations required are contractions (C) and expansions (E) :

$$\textbf{(C)} \quad \mathbf{a}^{\ell}\ \mathbf{a} \to \mathbf{1}\ ,\ \ \mathbf{a}\ \mathbf{a}^{r} \to \mathbf{1}\ ,$$
$$\textbf{(E)} \quad \mathbf{1} \to \mathbf{a}\ \mathbf{a}^{\ell}\ ,\ \ \mathbf{1} \to \mathbf{a}^{r}\ \mathbf{a}\ ,$$

5. Computations and Short Term Memory

In (2005) Chomsky argues that the primitive principles of grammar must also provide what George Miller called chunking identification of phonological words in the string of phonetic units. In LSLT (p. 165), he adopted Harris (1955) proposal, in a different framework, for identifying morphemes in terms of transitional probabilities, though morphemes do not have the required beads-on-a-string property. The basic problem, as noted in LSLT, is to show that such statistical methods of chunking can work with a realistic corpus.

In LSLT, it was assumed that the next step would be assignment of chunked items to syntactic categories, again by general principles of data analysis. A proposal with an information-theoretic flavor was tried by hand calculations in that precomputer age, with suggestive results, but the matter has never been pursued. The assumption of LSLT was that higher levels of linguistic description, including morphemes, are determined by a general format for rule systems provided by UG.

N. Chomsky, *Three Factors in Language Design*, **Linguistic Inquiry, 2005**
N. Chomsky, 1955, *Logical Structure of Linguistic Theory* **(LSLT), Univ. of Chicago Press, 1985.**

6. Language

The free pregroup for a language is generated by a partially ordered set of basic types such as:

s, $\tilde{s}$, $\overline{s}$, σ	**declarative sentence, sentential complements ,**
q, $\overline{q}$	**questions, indirect questions ,**
i, $\tilde{i}$, $\overline{i}$, i^*, j, $\overline{j}$, $\overline{\overline{i}}$	**infinitives, infinitival complements ,**
π	**subject ,**
o	**direct object ,**
ω	**indirect object ,**
λ	**locative phrase ,**
τ	**temporal phrase .**

6.1. Lexicon

The types o, ω, λ, τ are assigned to the arguments subcategorized by a verb; the types $\tilde{s}$, $\overline{s}$ indicate super types of the basic type of sentences s;

The type $\bar{\bar{i}}$ represents the maximal projection of a variety of infinitival complements i, $\tilde{i}$, $\bar{i}$, i^*, j, $\bar{j}$

(1) (*to run*) **correre :** i , $i\,\lambda^{\ell}$

(2) (*to see*) **vedere :** i , $i\,o^{\ell}$

(3) (*to give*) **dare :** $i\,\omega^{\ell}o^{\ell}$, $i\,o^{\ell}\omega^{\ell}$

6.2. Features

Numerical indices are assigned to simple types to express feature information such as: *person*, *gender*, *number*, *case* **for nouns, noun phrases and adjectives,** *tense*, *person* **for verbs,** ....

6.3. Templates

Featural information can be summarized in apposite matrices or templates, associated to any given inflectional type; e.g. for Latin the following have been introduced with respect to verbal inflection $\mathrm{C}^{\mathrm{a}}_{\mathrm{tp}}(\mathrm{V})$, adjectival inflection $\mathrm{D}^{\mathrm{g}}_{\mathrm{nc}}(\mathrm{A})$, and nominal inflection $\mathrm{D}^{\mathrm{g}}_{\mathrm{nc}}(\mathrm{N})$

$$\mathrm{C}^{\mathrm{a}}_{\mathrm{tp}}(\mathrm{V}) = \left[\begin{array}{l|l|l} a = ASPECT & t = TENSE & p = PERSON \\ \hline Basic : 1 & pres.ind. : 1 & first : 1 \\ Perfect : 2 & pres.subj. : 2 & second : 2 \\ Passive : 3 & imperf.ind. : 3 & third : 3 \\ & imperf.subj. : 4 & fourth : 4 \\ & future : 5 & fifth : 5 \\ & & sixth : 6 \end{array}\right]$$

7. Binding

We assume here a rather general notion of binding that includes all the linguistic domains in which an element (a variable) is bound by an antecedent. Following the studies in generative linguistics (see Chomsky 2005) we consider as instances of binding the domainns requiring conditions such as: *A-over-A Principle*, conditions on wh-extraction from wh-phrases (relatives and interrogatives), simplification of T-markers to base recursion (C. Fillmore), cyclicity principles (e.g. R. Freidin (1978), H. Lasnik 2007), taxonomy of islands (J. R. Ross (1967), M. Steedman (1985)), conditions on quantifiers scope (e.g. R. May (1980)), etc.

7.1. Wh-pronouns

Let us consider an interrogative clause involving the transitive verb incontrare (to meet), with type ($i\,\mathbf{o}^{\ell}$) and past participle ($p_2\,\mathbf{o}^{\ell}$) :

Chi ha incontrato Piero?
$(\overline{q}\ \hat{\pi}_3^{\ell\ell}\, q^{\,\ell})\ (q_1\ \pi_3^{\ell}\ p_2^{\ell})\ (p_2\, \mathbf{o}^{\ell})\ \mathbf{o} \rightarrow \overline{q}$

Maria si chiede chi ha incontrato Piero
$\pi_3\ (\pi_3^{r}\ \mathbf{s}_1\ \sigma^{\ell})\ (\overline{q}\ \hat{\pi}_3^{\ell\ell} q^{\,\ell})\ (q_1\ \pi_3^{\ell}\ p_2^{\ell})\ (p_2\ \mathbf{o}^{\ell})\ \mathbf{o} \rightarrow \mathbf{s}_1$

Chi dice che ha incontrato Piero?
$(\overline{q}\ \hat{\pi}_3^{\ell\ell} q^{\,\ell})\ (q_1\ \pi_3^{\ell}\ \sigma^{\ell})\ (\overline{\mathrm{s}}\ \tilde{\mathrm{s}}^{\ell})\ (\mathbf{s}_1\, p_2^{\ell})\ (p_2\, \mathbf{o}^{\ell})\ \mathbf{o} \rightarrow \overline{q}$

First example: the subject wh-pronoun Chi combines with the interrogative verb ha incontrato and the object Piero.

Second example: the corresponding indirect interrogative Maria si chiede chi ha incontrato Piero is derived, given $\overline{q} \rightarrow \sigma$.

Third example: the subject of the subordinate clause is omitted and we have TWO interpretations:

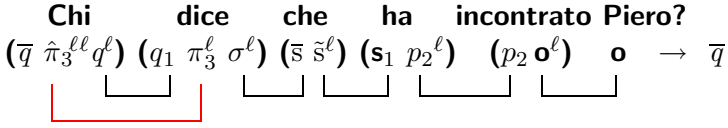
(I) There is somebody$_i$ saying that he$_i$ has met Piero

(II) There is somebody$_j$ that Piero$_i$ says he$_i$ has met

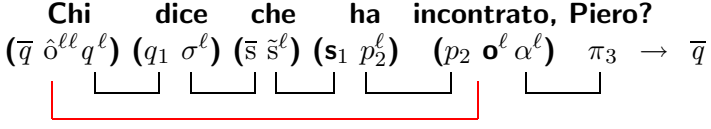
in (I) the wh-pronoun Chi is the SUBJECT, while in (II) Chi is the direct OBJECT.

We so represent the two interpretations:

Interpretation (I) :



Interpretation (II) :



Under the second interpretation, the wh-pronoun Chi takes type ($\overline{q}\ \hat{\mathrm{o}}^{\ell\ell}\, q^{\ell}$) and type ($p_2\ \mathbf{o}^{\ell}\, \alpha^{\ell}$) is assigned to the past participle incontrato, where $\pi_3 \to \alpha$

The link connecting the types ($\hat{\mathrm{o}}^{\ell\ell}\, \mathbf{o}^{\ell}$) shows that a long distance dependency is created between the wh-pronoun Chi and the direct object of the verb incontrare in the subordinate clause. An emphatic interpretation with the subject Piero in sentence final position.

We can add further information to the context, e.g. a new subject Maria and a person such that he has met Piero: "Maria is wondering who says he has met Piero"; we have:

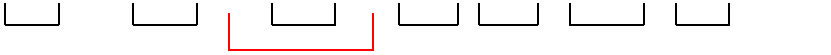
(I) Maria is wondering if there is somebody$_i$ saying that he$_i$ has met Piero$_j$

(II) Maria is wondering if there is somebody$_i$ whom$_i$ Piero$_j$ says he$_j$ has met him$_i$.

(Maria is wondering who says that he has met Piero)

Maria si chiede chi dice che ha incontrato Piero

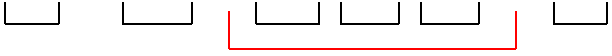
$\pi_3\ (\pi_3^r\ \mathbf{s}_1\ \sigma^\ell)\ (\overline{q}\ \hat{\pi}_3^{\ell\ell}\ q^\ell)\ (q\ \pi_3^\ell\ \sigma^\ell)\ (\overline{\mathrm{s}}\ \tilde{\mathrm{s}}^\ell)\ (\mathbf{s}_1\ p_2{}^\ell)\ (p_2\ \mathbf{o}^\ell)\ \mathbf{o} \rightarrow \mathbf{s}_1$



(Maria is wondering whom Piero says that he has met)

Maria si chiede chi dice che ha incontrato, Piero

$\pi_3\ (\pi_3^r\ \mathbf{s}_1\ \sigma^\ell)\ (\overline{q}\ \hat{\mathrm{o}}^{\ell\ell}\ q^\ell)\ (q\ \sigma^\ell)\ (\overline{\mathrm{s}}\ \tilde{\mathrm{s}}^\ell)\ (\mathbf{s}_1\ \mathbf{o}^\ell\ \alpha^\ell)\ \pi_3 \rightarrow \mathbf{s}_1$



A simpler analysis results if we omit the subject of the subordinate clause:

(Maria is wondering whom do you say to have met)

Maria si chiede chi dici che hai incontrato

$\pi_3\ (\pi_3^r\ \mathbf{s}_1\ \sigma^\ell)\ (\overline{q}\ \hat{\mathrm{o}}^{\ell\ell}\ q^\ell)\ (q_1\ (\pi_2^\ell)\ \sigma^\ell)\ (\overline{\mathrm{s}}\ \tilde{\mathrm{s}}^\ell)\ (\mathbf{s}_1\ \mathbf{o}^\ell) \rightarrow \mathbf{s}_1$

$= \pi_3\ (\pi_3^r\ \mathbf{s}_1\ \sigma^\ell)\ (\overline{q}\ \hat{\mathrm{o}}^{\ell\ell}\ q^\ell)\ (q_1\ \sigma^\ell)\ (\overline{\mathrm{s}}\ \tilde{\mathrm{s}}^\ell)\ (\mathbf{s}_1\ \mathbf{o}^\ell) \rightarrow \mathbf{s}_1$

