Categories, functions, predicates, concepts and the argument structure of the sentence

Francisco J. Salguero-Lamillar
salguero@us.es

Grupo de Lógica, Lenguaje e Información
Universidad de Sevilla

Some previous words

- Natural Language is the most powerful tool for generating complex concepts
- The relation among concepts and words is very natural, but it is not a biunivocal relation
- In every human language we can have many-to-one as well as one-to-many relations among words and concepts
- On the other hand, it is possible in all languages expressing concepts by means of structural rules that use more than one single word

Some previous words

- Natural Language is the most powerful tool for generating complex concepts
- The relation among concepts and words is very natural, but it is not a biunivocal relation
- In every human language we can have many-to-one as well as one-to-many relations among words and concepts
- On the other hand, it is possible in all languages expressing concepts by means of structural rules that use more than one single word

Some previous words

- Natural Language is the most powerful tool for generating complex concepts
- The relation among concepts and words is very natural, but it is not a biunivocal relation
- In every human language we can have many-to-one as well as one-to-many relations among words and concepts
- On the other hand, it is possible in all languages expressing concepts by means of structural rules that use more than one single word

Some previous words

- Natural Language is the most powerful tool for generating complex concepts
- The relation among concepts and words is very natural, but it is not a biunivocal relation
- In every human language we can have many-to-one as well as one-to-many relations among words and concepts
- On the other hand, it is possible in all languages expressing concepts by means of structural rules that use more than one single word

Aim and purposes

- Philosophers, logicians, linguists and even mathematicians have tried to decipher the mechanisms by which concepts are constructed from the meaning of words
- One way to achieve this was to study lexical meaning and its combinatorial properties
- Our purpose is to explore the seminal ideas that have resulted in categorial grammars and their relationship with other grammatical models and actual theories of meaning

Aim and purposes

- Philosophers, logicians, linguists and even mathematicians have tried to decipher the mechanisms by which concepts are constructed from the meaning of words
- One way to achieve this was to study lexical meaning and its combinatorial properties
- Our purpose is to explore the seminal ideas that have resulted in categorial grammars and their relationship with other grammatical models and actual theories of meaning

Aim and purposes

- Philosophers, logicians, linguists and even mathematicians have tried to decipher the mechanisms by which concepts are constructed from the meaning of words
- One way to achieve this was to study lexical meaning and its combinatorial properties
- Our purpose is to explore the seminal ideas that have resulted in categorial grammars and their relationship with other grammatical models and actual theories of meaning

The phenomenological origin of Categorial Grammar

"The task of an accomplished science of meanings would be to investigate the law-governed, essence-bound structure of meanings and the laws of combination and modification of meaning which depend upon these, also to reduce such laws to the least number of independent elementary laws. We should obviously also need to track down the primitive meaning-patterns and their inner structures, and, in connection with these, to fix the pure categories of meaning which circumscribe the sense and range of the indeterminates —the 'variables' in a sense exactly analogous to that of mathematics— that occur in such laws." (Husserl, *Fourth Logical Investigation*, § 13, 68)

Antecedents

- Aristotle: *The Categories*
  - Univocal, equivocal and derivative words
  - Predicates and predication (sentence structure): what is said *of* a subject and what is *in* a subject (an inherent concept)
- Medieval Speculative Grammar:
  - *Modi essendi* (the way to exist)
  - *Modi intelligendi* (the way to be conceptualized)
  - *Modi significandi* (the way to mean)

Antecedents

- Aristotle: *The Categories*
  - Univocal, equivocal and derivative words
  - Predicates and predication (sentence structure): what is said *of* a subject and what is *in* a subject (an inherent concept)
- Medieval Speculative Grammar:
  - *Modi essendi* (the way to exist)
  - *Modi intelligendi* (the way to be conceptualized)
  - *Modi significandi* (the way to mean)

Antecedents

- Rational Grammar:
  - J. C. Scaligero: *De causis linguae latinae libri XIII* (1540)
  - El Brocense: *Minerva sive de causis linguae latinae* (1587)
  - *Port-Royal Grammar* (1660)
- G. W. Leibniz: *De arte combinatoria* (1666) and the project of *mathesis universalis*
  - Goal: a universal science based on symbols
  - Method: linguistic analysis from complex terms to their most simple "formal parts" (indefinable terms)
  - Tools: these formal parts are represented by mathematical symbols and a few rules for their combination must be given
  - Indefinable mathematical symbols and the rules of combination describe a universal logic of meaning and discovery

Antecedents

- Rational Grammar:
  - J. C. Scaligero: *De causis linguae latinae libri XIII* (1540)
  - El Brocense: *Minerva sive de causis linguae latinae* (1587)
  - *Port-Royal Grammar* (1660)
- G. W. Leibniz: *De arte combinatoria* (1666) and the project of *mathesis universalis*
  - Goal: a universal science based on symbols
  - Method: linguistic analysis from complex terms to their most simple "formal parts" (indefinable terms)
  - Tools: these formal parts are represented by mathematical symbols and a few rules for their combination must be given
  - Indefinable mathematical symbols and the rules of combination describe a universal logic of meaning and discovery

Husserl's general science of significations

- E. Husserl: *Fourth Logical Investigation* (1900)
- He proposes a concept of grammar based on *a priori* laws that determine linguistic meaning rather than "exclusively on psychology and other empirical disciplines".
- That is to say, he returns to "the old idea of a general grammar" instead of the new empiricist trends of the century
- This is a semantic perspective, so that linguistic expressions are for him significations (*Bedeutungen*) that are assigned semantic categories (*Bedeutungskategorien*)
- These significations can be simple (the lexicon) and compound (sentences), so that simple significations are only partial meanings that require completion to give full meanings by means of certain combinatorial rules

Husserl's general science of significations

- E. Husserl: *Fourth Logical Investigation* (1900)
- He proposes a concept of grammar based on *a priori* laws that determine linguistic meaning rather than "exclusively on psychology and other empirical disciplines".
- That is to say, he returns to "the old idea of a general grammar" instead of the new empiricist trends of the century
- This is a semantic perspective, so that linguistic expressions are for him significations (*Bedeutungen*) that are assigned semantic categories (*Bedeutungskategorien*)
- These significations can be simple (the lexicon) and compound (sentences), so that simple significations are only partial meanings that require completion to give full meanings by means of certain combinatorial rules

Husserl's general science of significations

- E. Husserl: *Fourth Logical Investigation* (1900)
- He proposes a concept of grammar based on *a priori* laws that determine linguistic meaning rather than "exclusively on psychology and other empirical disciplines".
- That is to say, he returns to "the old idea of a general grammar" instead of the new empiricist trends of the century
- This is a semantic perspective, so that linguistic expressions are for him significations (*Bedeutungen*) that are assigned semantic categories (*Bedeutungskategorien*)
- These significations can be simple (the lexicon) and compound (sentences), so that simple significations are only partial meanings that require completion to give full meanings by means of certain combinatorial rules

Husserl's general science of significations

- E. Husserl: *Fourth Logical Investigation* (1900)
- He proposes a concept of grammar based on *a priori* laws that determine linguistic meaning rather than "exclusively on psychology and other empirical disciplines".
- That is to say, he returns to "the old idea of a general grammar" instead of the new empiricist trends of the century
- This is a semantic perspective, so that linguistic expressions are for him significations (*Bedeutungen*) that are assigned semantic categories (*Bedeutungskategorien*)
- These significations can be simple (the lexicon) and compound (sentences), so that simple significations are only partial meanings that require completion to give full meanings by means of certain combinatorial rules

Husserl's general science of significations

- E. Husserl: *Fourth Logical Investigation* (1900)
- He proposes a concept of grammar based on *a priori* laws that determine linguistic meaning rather than "exclusively on psychology and other empirical disciplines".
- That is to say, he returns to "the old idea of a general grammar" instead of the new empiricist trends of the century
- This is a semantic perspective, so that linguistic expressions are for him significations (*Bedeutungen*) that are assigned semantic categories (*Bedeutungskategorien*)
- These significations can be simple (the lexicon) and compound (sentences), so that simple significations are only partial meanings that require completion to give full meanings by means of certain combinatorial rules

Comparing with Frege's Principles

- *Principle of Compositionality*: The meaning of a complex expression is a function on the meaning of the most simple expressions that compound it and the rules of combination used over these simple expressions to generate the complex one
- *Context Principle*: It is necessary to consider the words as a part of the sentence when we ask for their meaning [...] It is enough when a whole complete sentence has a meaning; thereby also its parts receive their content

Categorial Grammar is a Logical Grammar

- However, Husserl's proposal is not logicist; i.e.: there is no identification between the laws of logic and the rules of grammar
- Nevertheless, for him the rules of grammar are logical; i.e.: there is a (certain kind of) logic behind the combinatorial rules of significations
- Every language works on the basis of a general logic that establishes laws of possibility and exclusion, and these laws —and not others— are making up what he calls "pure grammar"
- So pure logical grammar is conceived as a set of analytical laws common to all languages

Categorial Grammar is a Logical Grammar

- However, Husserl's proposal is not logicist; i.e.: there is no identification between the laws of logic and the rules of grammar
- Nevertheless, for him the rules of grammar are logical; i.e.: there is a (certain kind of) logic behind the combinatorial rules of significations
- Every language works on the basis of a general logic that establishes laws of possibility and exclusion, and these laws —and not others— are making up what he calls "pure grammar"
- So pure logical grammar is conceived as a set of analytical laws common to all languages

Categorial Grammar is a Logical Grammar

- However, Husserl's proposal is not logicist; i.e.: there is no identification between the laws of logic and the rules of grammar
- Nevertheless, for him the rules of grammar are logical; i.e.: there is a (certain kind of) logic behind the combinatorial rules of significations
- Every language works on the basis of a general logic that establishes laws of possibility and exclusion, and these laws —and not others— are making up what he calls "pure grammar"
- So pure logical grammar is conceived as a set of analytical laws common to all languages

Categorial Grammar is a Logical Grammar

- However, Husserl's proposal is not logicist; i.e.: there is no identification between the laws of logic and the rules of grammar
- Nevertheless, for him the rules of grammar are logical; i.e.: there is a (certain kind of) logic behind the combinatorial rules of significations
- Every language works on the basis of a general logic that establishes laws of possibility and exclusion, and these laws —and not others— are making up what he calls "pure grammar"
- So pure logical grammar is conceived as a set of analytical laws common to all languages

Categories like functions

- Husserl proposes a rule-based grammar over semantic connections through which it is possible to integrate the significant incomplete parts with the right parts to complete them (in a similar sense to the saturation of a function by its arguments, according to Frege)
- It is evident, therefore, that the mode of composition depends on the set of categories of significance chosen, observing a certain set of universal principles (Casadio 1988):
  - Any linguistic expression must belong to a category of significance
  - Any meaningful expression is the result of the integration of its parts, depending the integration mode on the categories of significance to which each part belongs
  - By replacing a part of a meaningful expression by an expression of a different category of significance, the first ever becomes non-meaningful

Categories like functions

- Husserl proposes a rule-based grammar over semantic connections through which it is possible to integrate the significant incomplete parts with the right parts to complete them (in a similar sense to the saturation of a function by its arguments, according to Frege)

- It is evident, therefore, that the mode of composition depends on the set of categories of significance chosen, observing a certain set of universal principles (Casadio 1988):
    - Any linguistic expression must belong to a category of significance
    - Any meaningful expression is the result of the integration of its parts, depending the integration mode on the categories of significance to which each part belongs
    - By replacing a part of a meaningful expression by an expression of a different category of significance, the first ever becomes non-meaningful

The Lwów-Warsaw Circle

- The first formal applications of Husserl's proposal treat complex categories like functions that formalize predication and other syntactic connections: Lesniewki (1927-1931) and Ajdukiewicz (1935)
- Ajdukiewicz categorial grammar is based on two basic categories (n, s) and two functors that define two different function types $A \backslash B$ and B/A
- Every expression of a language —simple or complex— belongs to a basic category or to a functional type defined over the two basic categories

The Lwów-Warsaw Circle

- The first formal applications of Husserl's proposal treat complex categories like functions that formalize predication and other syntactic connections: Lesniewki (1927-1931) and Ajdukiewicz (1935)
- Ajdukiewicz categorial grammar is based on two basic categories (n, s) and two functors that define two different function types $A \backslash B$ and B/A
- Every expression of a language —simple or complex— belongs to a basic category or to a functional type defined over the two basic categories

The Lwów-Warsaw Circle

- The first formal applications of Husserl's proposal treat complex categories like functions that formalize predication and other syntactic connections: Lesniewki (1927-1931) and Ajdukiewicz (1935)
- Ajdukiewicz categorial grammar is based on two basic categories (n, s) and two functors that define two different function types $A\backslash B$ and B/A
- Every expression of a language —simple or complex— belongs to a basic category or to a functional type defined over the two basic categories

AB Categorial Grammar

- Category n is the category of those expressions that refer to an individual
- Category s is the category of those expresions that refer to a proposition
- The function type $A\backslash B$ is interpreted as the type of an expresion that results of type B when it is preceded by an expression of type A
- The function type B/A is interpreted as the type of an expresion that results of type B when it is followed by an expression of type A

AB Categorial Grammar

- Category n is the category of those expressions that refer to an individual
- Category s is the category of those expresions that refer to a proposition
- The function type $A \backslash B$ is interpreted as the type of an expresion that results of type B when it is preceded by an expression of type A
- The function type B/A is interpreted as the type of an expresion that results of type B when it is followed by an expression of type A

AB Categorial Grammar

- Category n is the category of those expressions that refer to an individual
- Category s is the category of those expresions that refer to a proposition
- The function type $A\backslash B$ is interpreted as the type of an expresion that results of type B when it is preceded by an expression of type A
- The function type B/A is interpreted as the type of an expresion that results of type B when it is followed by an expression of type A

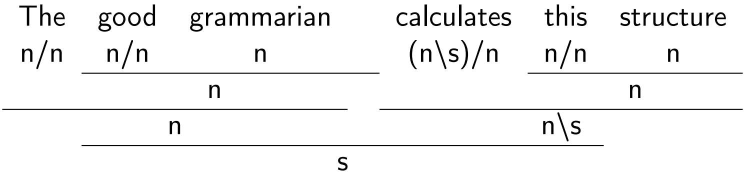
AB Categorial Grammar

- Category n is the category of those expressions that refer to an individual
- Category s is the category of those expresions that refer to a proposition
- The function type $A\backslash B$ is interpreted as the type of an expresion that results of type B when it is preceded by an expression of type A
- The function type B/A is interpreted as the type of an expresion that results of type B when it is followed by an expression of type A

An algebraic grammatical calculus

- Following Lambek (1958), AB Grammar can be defined as a kind of algebraic calculus
- AB Grammar is a tuple $\langle \Sigma, Prim, Tp, \triangleright \rangle$, where:
  - Σ is a finite set of symbols
  - $Prim$ is the set of primitive types
  - $Tp(Prim)$ is the set of all types built over the set of primitive types such that it is the smallest set that satisfies $Prim \subseteq Tp(Prim)$ and if $X, Y \in Tp(Prim)$ then $(X/Y), (Y \backslash X) \in Tp(Prim)$
  - $\triangleright$ is a relation that assignes a lexical element to a categorial type such that $(\triangleright) \subseteq Tp(Prim) \times \Sigma$

An example



The parser is ((the (good grammarian)) (calculates (this structure))), corresponding to the following phrase structure:

$[_S\ [_{NP}[_{Det}\, \mathit{the}[_{Adj}\, \mathit{good}[_N\, \mathit{grammarian}]\ [_{VP}[_V\, \mathit{calculates}[_{NP}[_{Det}\, \mathit{this}[_N\, \mathit{structure}]\ ]$

An example

- We can see several characteristics of AB Ctegorial Grammar in this example:
  1. Every word is assigned a category or a categorial type
  2. Complex concepts as "good grammarian" or "this structure" are assigned a categorial type calculated as a function over the types of the simplest ones
  3. All the process is a kind of predication: we predicate "good" of "grammarian", "calculates" of "this structure", and "calculates this structure" of "the good grammarian"

Expanding to semantics

- AB Grammar is a syntactic calculus, but it can be expanded to semantics, reinterpreting categories and functional types
- That was the intention of Richard Montague when he proposed the semantic theory that bears his name (Montague 1973)
- Montague Grammar is based on the Categorial Grammar proposal, but it is semantically oriented:
  - Two basic categories: e and t (correspondig to the AB Grammar categories n and s)
  - Just one function to get categorial types, represented by ordered tuples: $\langle e, t\rangle \mid \langle e, \langle e, t\rangle\rangle \mid \langle\langle e, t\rangle, t\rangle \mid \langle t, t\rangle \mid \langle\langle e, t\rangle, \langle e, t\rangle\rangle \mid \langle e, \langle\langle e, t\rangle, \langle e, t\rangle\rangle\rangle$

Basic Lambek Calculus

- We can combine MG with a set of rules based on Lambek Sequent Calculus (just a hint):

$$\frac{A \qquad \langle A, B\rangle}{B} \qquad \frac{\langle A, B\rangle \qquad \langle B, C\rangle}{\langle A, C\rangle} \qquad \frac{A}{\langle\langle A, B\rangle, B\rangle}$$

$$\frac{\langle\langle\langle A, B\rangle, B\rangle, C\rangle}{\langle A, C\rangle} \qquad \frac{\langle B, C\rangle}{\langle\langle A, B\rangle, \langle A, C\rangle\rangle}$$

Correspondence with a (context free) Generalized Phrase Structure Grammar

- $S \rightarrow NP\ VP$: $\quad t \longrightarrow \dfrac{\langle\langle e,t\rangle,t\rangle \qquad \langle e,t\rangle}{t}$

- $NP \rightarrow PN$: $\quad \langle\langle e,t\rangle,t\rangle \longrightarrow \dfrac{e}{\langle\langle e,t\rangle,t\rangle}$

- $NP \rightarrow Det\ CN$: $\quad \langle\langle e,t\rangle,t\rangle \longrightarrow \dfrac{\langle\langle e,t\rangle,\langle\langle e,t\rangle,t\rangle\rangle \qquad \langle e,t\rangle}{\langle\langle e,t\rangle,t\rangle}$

- $VP \rightarrow TV\ NP$: $\quad \langle e,t\rangle \longrightarrow \dfrac{\langle e,\langle e,t\rangle\rangle \qquad \langle\langle e,t\rangle,t\rangle}{\langle e,t\rangle}$

Limits of Categorial Grammars

- The most important objection to AB Categorial Grammars is that they are equivalent to context-free phrase structure grammars, which makes them not suitable for the description of certain syntactic phenomena
- Nevertheless, these grammars are useful for the semantic description of phenomena such as quantifier scope, anaphoric relations, ambiguities between *de dicto* and *de re* interpretations of certain terms or the distinction between extensional verbs and intensional verbs like propositional attitudes
- But it is possible to combine Montague Grammar with a Dynamic Predicate Logic (Groenendijk & Stokhof 1989, 1991) as well as to use Categorial Grammar as a basis for a Unification Grammar based on types

Unification and argument structure

- Basic categories *n* and *s* can be defined as feature structures. For example, the feature structure for nouns:

$$\begin{bmatrix} \mathit{CONT} & +, - \\ \mathit{DEF} & +, - \\ \mathit{CASE} & \mathit{NOM}, \mathit{ACC}, \mathit{OBL} \\ \mathit{LEVEL} & 1, 2 \\ \mathit{ROLE} & \mathit{AG}, \mathit{PAT}, \mathit{GOAL}, \mathit{INSTR}, \mathit{DEST}, \mathit{LOC} \\ \mathit{CONC} & \\ \begin{bmatrix} \mathit{NUM} & \mathit{SING}, \mathit{PL} \\ \mathit{GEN} & \mathit{MASC}, \mathit{FEM}, \mathit{NEUT} \end{bmatrix} & \end{bmatrix}$$

Unification and argument structure

- When we treat categories as feature structures, we are opening a door to apply the rules of a categorial calculus on a basic argument structure of the sentence with unification of features
- The argument structure of a sentence is a basic predicate structure that supports recursivity (i.e.: all its arguments can be replaced by complete argument structures):

$$MOD((PRED(arg_2),(arg_3))arg_1)SAT$$

- In an argument structure, only the predicate *PRED* and the external argument arg_1 are necessary. The inner arguments arg_2 and arg_3 are not necessary
- *MOD* is any kind of modality
- *SAT* might be a complementary squence of terms or structures

Conclusion

- Categories can be seen as functions that relate a (complex) meaning and certain combinatorial properties to get more complex categories
- This is a way of representing concepts as the result of predication
- The different kind of predicative relations can be reduced to a few categories and a few rules for getting categorial types
- Categories and categorial types can be treated as feature structures that participate in recursive argument structures built over predicates and arguments

Conclusion

- Categories can be seen as functions that relate a (complex) meaning and certain combinatorial properties to get more complex categories
- This is a way of representing concepts as the result of predication
- The different kind of predicative relations can be reduced to a few categories and a few rules for getting categorial types
- Categories and categorial types can be treated as feature structures that participate in recursive argument structures built over predicates and arguments

Conclusion

- Categories can be seen as functions that relate a (complex) meaning and certain combinatorial properties to get more complex categories
- This is a way of representing concepts as the result of predication
- The different kind of predicative relations can be reduced to a few categories and a few rules for getting categorial types
- Categories and categorial types can be treated as feature structures that participate in recursive argument structures built over predicates and arguments

Conclusion

- Categories can be seen as functions that relate a (complex) meaning and certain combinatorial properties to get more complex categories
- This is a way of representing concepts as the result of predication
- The different kind of predicative relations can be reduced to a few categories and a few rules for getting categorial types
- Categories and categorial types can be treated as feature structures that participate in recursive argument structures built over predicates and arguments

References

1. Ajdukiewicz, K. (1935): "Die syntaktische Konnexität". *Studia Philosophica* 1: 1–27.
2. Bar-Hillel, Y. (1950): "On syntactical categories". *The Journal of Symbolic Logic* 15: 1–16
3. Bar-Hillel, Y. (1953) "A quasi-arithmetical notation for syntactic description". *Language* XXIX: 47–58.
4. Groenendijk, J. & M. Stokhof (1989): "Dynamic Montague Grammar". In *Papers from the Second Symposium on Logic and Language*. Akademiai Kiadoo, 3–48.
5. Groenendijk, J. & M. Stokhof (1991): "Dynamic predicate logic". *Linguistics and Philosophy*, 14: 39–100.
6. Husserl, E. (1990): *Logische Untersuchungen*. Niemeyer.
7. Lambek, J. (1958): "The mathematics of sentence structure". *American Mathematical Monthly* 65(3): 154–170.
8. Montague, R. (1973): "The Proper Treatment of Quantification in Ordinary English". In Jaakko Hintikka, Julius Moravcsik, Patrick Suppes (eds.): *Approaches to Natural Language*. Dordrecht: Reidel, 221–242.

THANK YOU VERY MUCH

MUITO OBRIGADO

MUCHAS GRACIAS