



HAL
open science

O Léxico-Gramática, um método de descrição linguística

Eric Laporte

► **To cite this version:**

Eric Laporte. O Léxico-Gramática, um método de descrição linguística. 2013, pp.1-125. hal-00823401

HAL Id: hal-00823401

<https://hal.science/hal-00823401v1>

Submitted on 16 May 2013

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

O Léxico-Gramática, um método de descrição linguística

Éric Laporte

Universidade Federal do Espírito Santo
Université Paris-Est, LIGM (UMR 8049 CNRS)

<http://igm.univ-mlv.fr/~laporte/>

Sumário

Objetivos históricos

Princípios

Modo de formalização

Resultados

Objetivos históricos

Entender a sintaxe e da semântica

Rigor científico

A linguística deve ser tão rigorosa quanto as outras ciências

Esquadrinhamento do léxico

Imprescindível para observação dos fatos linguísticos

Nenhum objetivo relativo a mecanismos mentais

Rigor científico

O problema da intuição semântica

A observação semântica se baseia na intuição

A intuição é diferente de um locutor para outro

Moleque é pejorativo?

Respaldo da intuição semântica

Teste operatório: teste aplicável e eficaz

Utiliza a realidade diretamente observável

Rigor científico

O problema da intuição semântica

A observação semântica se baseia na intuição

A intuição é diferente de um locutor para outro

Moleque é pejorativo?

Respaldo da intuição semântica

Teste operatório: teste aplicável e eficaz

Utiliza a realidade diretamente observável

Seu moleque!

Distanciamento com semântica

Origina-se com Sapir e Bloomfield

Linguistas americanos

Estudo dos idiomas índios por linguistas de cultura europeia

Falta de intuição semântica

Distribucionalismo

Investigação dos paradigmas

Distanciamento com semântica

Origina-se com Sapir e Bloomfield

Linguistas americanos

Estudo dos idiomas índios por linguistas de cultura europeia

Falta de intuição semântica

Distribucionalismo

Investigação dos paradigmas

Seu moleque!

?* *Sua estrela!*

Seu chulezudo!

Moleque querido!

Estrela querida!

* *Chulezudo querido!*

Distanciamento com semântica

?* *Sua estrela!*

Estrela querida!

Seu chulezudo!

* *Chulezudo querido!*

Esse método tornou a linguística mais científica:

- uma forma de experimentação
- uso de informações negativas, como em outras ciências

Antes, a linguística fazia parte da literatura

A análise distribucional é uma ferramenta metodológica fundamental

Bibliografia: Bloomfield (1933)

Distanciamento com semântica

A semântica é um objetivo central da descrição linguística

A observação semântica necessita métodos indiretos

O distribucionalismo é uma teoria, mas trata sobretudo de **métodos**

Métodos aplicados também ao idioma nativo do linguista

Zellig Harris

Linguista americano, distribucionalista, formado em matemática

Dúvida metódica

Observação reprodutível

Que pode ser repetida por outros observadores, com os mesmos resultados

Um termo das ciências experimentais

Reprodutibilidade

Uma exigência teórica e também prática

Ponto de vista metodológico

A construção de recursos linguísticos é um trabalho de equipe

Necessita acordo entre especialistas

Léxico-gramática dos verbos franceses:

encontros semanais de 6 pessoas durante 18 anos

Ponto de vista comercial

O resultado deve convencer os usuários

Reprodutibilidade

Cautela na observação semântica

A intuição de **aceitabilidade** varia menos

Associar cada propriedade semântica a um ou
vários fatos diretamente observáveis, através
da aceitabilidade

Resultados mais reprodutíveis

Esquadrinhamento do léxico

Imprescindível para observação linguística

Maurice Gross

Linguista francês, formado em engenharia de armas

Mecanismos regulares da linguagem

Interrogação, negação, flexão...: óbvios

Comportamentos caóticos

Diferenças entre entradas lexicais:
subestimadas

Esquadrinhamento do léxico

Imprescindível para observação linguística

Maurice Gross

Linguista francês, formado em engenharia de armas

Mecanismos regulares da linguagem

Interrogação, negação, flexão...: óbvios

Comportamentos caóticos

Diferenças entre entradas lexicais:
subestimadas

*João (é + *está) adolescente*

João (é + está) casado

*João (*é + está) morto*

Sumário

Objetivos históricos

Princípios

Modo de formalização

Resultados

Esquadrinhamento do léxico

Imprescindível por causa das diferenças inesperadas entre entradas lexicais
Construção de dicionários

Léxico-gramática do francês

13 900 entradas lexicais de verbos

Total: 76 000 entradas

Bibliografia: Gross (1975)

Contexto frasal

O léxico-gramática estuda palavras só dentro de **contextos de uma frase** pelo menos

jogar

Zé joga baralho

Zé joga a bola para fora do terreno

Um verbo com um sentido vem com uma construção sintática (ou várias)

Contexto frasal

A sintaxe da **frase elementar** já está longe de ser conhecida completamente, até em inglês
O quadro de estudo mais conveniente é a frase elementar (Gross, 1975)
« A unidade mínima de sentido é a frase elementar »

Contexto frasal

« A unidade mínima de sentido é a frase elementar »

Zé joga baralho

Segundo outras tradições, a unidade mínima de sentido pode ser:

- o morfema (morfologia): *jog-*
- a palavra (semântica lexical): *jogo*
- a frase sem sujeito (lexicografia): *jogar baralho*

Predicados semânticos e argumentos

Frases elementares correspondem a predicados semânticos e argumentos

Predicado semântico = núcleo do predicado

Zé joga baralho

predicado semântico: *jogar*

argumentos: *Zé, baralho*

Bibliografia: Tesnière, 1959; Montague, 1970

Relações entre construções

(1) *João anuncia que tem câncer*

= (2) *João anuncia ter câncer*

(3) *Ana imagina que está grávida*

= (4) *Ana imagina estar grávida*

A mesma relação entre construções ocorre entre (1) e (2) e entre (3) e (4)

Transformação sintática, o objeto central da sintaxe (Harris, 1964)

Transformacionalismo

Um adjetivo qualifica atitude,
opinião, aparência...?

João é cortês

atitude, comportamento

João é abolicionista

opinião, posição

João é desleixado

aparência

Associamos etiquetas semânticas por intuição

Um adjetivo qualifica atitude?

João é cortês

= João tem uma atitude cortês

O sentido é conservado

Fica mais preciso

Nuance aspectual

Substantivo apropriado

João é cortês

= João tem um(a) cortês

Qual é o substantivo com a maior plausibilidade de ocorrência?

Uma sequência é apropriada quando tem tanta plausibilidade de ocorrência que os falantes podem omiti-la sem perda de informação

Um adjetivo qualifica atitude,
opinião, aparência...?

João é abolicionista

= *João tem opiniões abolicionistas*

João é desleixado

= *João tem uma aparência desleixada*

O cartão é quadrado

= *O cartão tem uma forma quadrada*

Linguagem	Metalinguagem
<i>João é cortês</i>	atitude
<i>João é abolicionista</i>	opinião
<i>João é desleixado</i>	aparência
<i>O cartão é quadrado</i>	forma

Reduzir a distância entre linguagem e metalinguagem

Bibliografia: Harris (1964)

Realidade observável

Qualquer informação incluída na descrição deve ser verificável a partir da realidade diretamente observável

O complemento direto de *encher* é obrigatório?

João encheu o caminhão de móveis

**João encheu de móveis*

É obrigatório

A sintaxe é considerada uma **ciência experimental**

Qual é o sujeito implícito de uma infinitiva?

O prefeito anunciou *à imprensa* ter gostado da festa

Com *anunciar*, o sujeito ausente da infinitiva é o **sujeito** da frase principal

A lei proíbe *a réus* ser candidatos

Com *proibir*, é o **complemento** da frase principal

Qual é o sujeito implícito de uma infinitiva?

E com *prometer*?

João prometeu *ao filho* ir à praia

Utilizamos a intuição para responder

Para utilizar o julgamento de aceitabilidade em vez da intuição, preparar um **teste**

Testar o sujeito implícito de uma infinitiva

Escolher sujeito e complemento pertencendo a **números** (sg, pl) diferentes e **concordar** o infinitivo com o sujeito implícito

Ele prometeu **a vocês** (*ir + *irem*) à praia

Nós prometemos **a você** (*irmos + ir*) à praia

Experiência e prova em sintaxe

Pergunta: N_0 *anunciar a* N_2 [*V-inf W*] aceita
correferência entre N_2 e o sujeito de *V-inf*?

Experiência:

- forjar *Anuncio a eles que passaram na pr.*
 Anuncio a eles terem passado na pr.
- julgar **Anuncio a eles terem passado na pr.*

Resposta: não aceita

Análise distribucional

Exploração dos paradigmas

Análise distribucional de formas predicativas

Explorar as possibilidades de substituição de um constituinte

[O São Paulo] joga [com o Santa Cruz]
[N] jogar [com N]

[O São Paulo] pode ser substituído por um sintagma nominal denotando pessoas

Restrições distribucionais

- * *A bola joga com o Santa Cruz*
- * *A fé joga com o Santa Cruz*
- * *Torcer joga com o Santa Cruz*

[O São Paulo] só pode ser substituído por um sintagma nominal humano

Restrições distribucionais

[As festas] incomodam [os vizinhos]

O bebê incomoda os vizinhos

A incerteza incomoda os vizinhos

Que eu toque bateria incomoda os vizinhos

[As festas] pode ser substituído por qualquer sintagma nominal ou oração completiva

Análise distribucional

Essa análise deve distinguir as interpretações

[Zé] joga [a bola] [no chão]

[Zé] joga [baralho]

Restrições distribucionais

[O prefeito] anuncia [à imprensa] [que é candidato]

[N] anuncia [a N] [que F]

que é candidato pode ser substituído por
oração infinitiva

O prefeito anuncia à imprensa ser candidato

Análise distribucional

A análise distribucional é aplicável a elementos que não são constituintes

[Zé] joga [a bola] [na caixa]

A preposição *em* pode ser substituída por *até, para, sob, sobre, entre, rumo a...*
dentro de, perto de, longe de...

embaixo de, em cima de, em torno de...
para dentro de, para fora de, para baixo de,
para cima de...

Tipos de substantivos

Humanos: *agência, bebê, candidato, gente, imprensa, namorada, Pelé, prefeito, São Paulo, show, turista, vizinho*

Concretos: *bar, bateria, bola, caixa, chão, cidade, hotel, quarto, sapato*

Abstratos: *fé, festa, futebol, imprensa, incerteza, jogo, voz*

Bibliografia: Guillet, 1986

Comparação semântica

Avaliação semântica absoluta

Ana é brasileira qualifica modo de vida?

Comparação semântica

(5) *Ana é brasileira*

(6) *O fato de morar ali abrasileira Ana*

A frase (6) denota uma mudança de estado para (5) e especifica a causa?

A comparação semântica é mais reprodutível do que a avaliação semântica absoluta (Gross, 1975)

Regularidade

Ana é brasileira/O fato de morar ali a brasileira

Ana

Ana é francesa/O fato de morar ali a francesa

Ana

Ana é portuguesa/O fato de morar ali

a portuguesa Ana

N_0 ser Adj/ N_{caus} a-Adj-ar N_0

Reprodução dos mesmos fatos:

As duas frases não têm o mesmo sentido

A **diferença de sentido é a mesma** com cada adjetivo étnico

Comparação semântica do 2º grau

Comparação do primeiro grau: entre formas

Comparação do segundo grau: entre
comparações do primeiro grau

A comparação do 2º grau é fácil e reprodutível

Um elemento essencial da noção de
transformação sintática

Princípios

Esquadrinhamento do léxico

Contexto frasal

Relações entre construções

Realidade observável

Análise distribucional

Comparação semântica

Sumário

Objetivos históricos

Princípios

Modo de formalização

Resultados

O que é formalizar?

Respeitar regras de forma pelas quais o conteúdo se torna mais preciso, mais nítido, mais claro

Exemplo

fazer cara de nojo/mau/feia

O que é formalizar?

Respeitar regras de forma pelas quais o conteúdo se torna mais preciso, mais nítido, mais claro

Exemplo

fazer cara de nojo/mau/feia

mais formalizado:

fazer cara (de (nojo + mau) + feia)

O que é um modelo?

Modelo informático: visão simplificada que permite um tratamento

Exemplo: modelo dos empregados de uma empresa para um sistema de gestão de salário

Informações representadas: nome, endereço, função

Informações negligenciadas: estilo de música preferido, cor dos cabelos, número de irmãos...

Modelo do léxico-gramática

Entradas lexicais

Construções “sintáticas”

Fenômenos não codificados

Estilo

Etimologia

Fatos dialetais (exceto com descrições separadas)

Fatos socioletais (exceto com descrições separadas)

Mecanismos mentais do falante

Sentido, realidade extra-linguística, conotações...: levados em conta mas não codificados

Entradas lexicais

As entradas lexicais são os elementos do léxico

Distinção de entradas lexicais

Na análise distribucional, distinguir as interpretações

[Zé] joga [a bola] [no chão]

[Zé] joga [baralho]

Muitas vezes acaba numa distinção de entradas lexicais

A distinção de entradas lexicais é importante porque a tradução em outras línguas pode ser diferente: *throw/play*

Distinção de entradas lexicais

Substituindo o conteúdo de um argumento, o sentido da frase muda

Primeiro caso: a diferença de sentido da frase resulta **logicamente** da diferença de sentido do argumento

Zé pegou a geladeira *Zé pegou o gato*

Zé pegou o arquivo *Zé pegou o dinheiro*

Não necessita distinção de entradas lexicais

Distinção de entradas lexicais

Segundo caso: a diferença de sentido do argumento produz na frase uma diferença de sentido **imprevisível**

Zé pegou a primeira geladeira

Zé pegou a primeira rua

Necessita distinção de entradas lexicais

Distinção de entradas lexicais

Casos duvidosos: a diferença de sentido produzida na frase é **previsível** ou **imprevisível** ?

João perdeu as chaves João perdeu Zé na festa
O carro pegou A moda pegou A cola pegou
A muda pegou O enxerto pegou

Distinção de entradas lexicais

*João perdeu as chaves João perdeu Zé na festa
O carro pegou A moda pegou A cola pegou
A muda pegou O enxerto pegou*

Distinção de entradas lexicais ou não?

É melhor recorrer a critérios sintáticos, que são
mais operatórios, confiáveis, reproduzíveis

Entrada lexical

Codificar as propriedades separadamente para cada sentido

O João jogou a bola para fora do terreno

N_0 V N_1 Loc N_2 sim

$N_1 =: Nconc$ sim

O João joga baralho

N_0 V N_1 Loc N_2 não

$N_1 =: Nconc$ não

Entrada lexical

Sem separação dos sentidos, a codificação das propriedades não faz sentido

Entrada lexical

Expressões fixas são entradas adicionais

O João jogou a bola para fora do terreno

N_0 V N_1 *Loc* N_2 sim

N_0 V N_1 sim

O João jogou fora o computador

N_0 V N_1 *Loc* N_2 não

N_0 V N_1 sim

Entrada lexical

Se codificamos uma palavra como ambígua entre duas entradas lexicais, deve existir pelo menos uma propriedade pela qual elas diferem

Geralmente a construção de base

O João jogou a bola para fora do terreno

N_0 V N_1 Loc N_2

O João jogou fora o computador N_0 V N_1

Fórmulas sintáticas

Representação codificada das construções sintáticas

seu N!

N₀ ser Adj

N₀ estar Adj

N₀ anunciar a N₂ [V-inf W]

Fórmulas sintáticas

Delimitar os constituintes

Fazer a análise distribucional dos constituintes

Substituir cada um por um símbolo (separando a preposição), levando em conta o resultado da análise

Substituir o verbo conjugado pelo infinitivo

[O São Paulo] joga [com o Santa Cruz]

[Nhum] jogar [(com + contra) Nhum]

Em geral, os colchetes são dispensáveis

Nhum jogar (com + contra) Nhum

Fórmulas sintáticas

*[Zé] fala [do hotel] [para o gerente]
Nhum falar de N para Nhum*

*[Zé] joga [baralho]
Nhum jogar Njogo*

*[O show] lotou [o hotel] [de turistas]
Nevento lotar Nconc de Nconc-pl*

Se necessário, numerar os constituintes

*N_0 evento lotar N_1 conc de N_2 conc-pl
= N_2 conc-pl lotar N_1 conc*

Aceitabilidade

Sequência aceitável

Sequência que é uma frase do idioma em estudo

João jogou a bola para fora do terreno

Sequência inaceitável

Sequência que não é

**João jogou baralho para fora do terreno*

O léxico-gramática considera a aceitabilidade como binária (sem intermediários), para simplificar

Uso linguístico

Uso linguístico: conjunto de hábitos de linguagem compartilhados por uma comunidade

O conhecimento do uso linguístico pode se adquirir

- por meio de formas atestadas,
- por meio da introspecção do falante nativo inserido na comunidade

Aceitabilidades negativas

A única fonte de aceitabilidades negativas é a **introspecção**

Um corpus de formas atestadas nunca traz tais informações, por causa da criatividade linguística

A introspecção necessita muito mais precauções, mas é a única forma de incluir **limites** na descrição

Informações negativas

O uso de informações negativas é essencial,
como em outras ciências

Popper (1959) define a noção de falseabilidade

Uma teoria científica, para comprovar
hipóteses, precisa fornecer:

- meios de comprovar que estão certas
- meios de comprovar que estão erradas

Aceitabilidade negativa cumpre esse papel

Informações negativas

Cientificamente, um fato negativo tem **pelo menos** a mesma importância de um positivo

A gramática gerativa não tem receio de usar

Pasteur (1857) demonstrou a existência de micróbios pela **não**-decomposição de um líquido selado

Bibliografia

Boons *et al.* (1976:29-52); Popper (1959)

Fixidez

Um conjunto de elementos linguísticos é fixo (ou cristalizado) se a manutenção do sentido da sequência necessita um bloqueio distribucional

*João **jogou fora** o computador*

≠ *João **jogou dentro** o computador*

≠ *João **arremessou fora** o computador*

O léxico-gramática considera a fixidez como binária (sem intermediários), para simplificar

Fixidez

Expressões fixas são entradas adicionais

O João jogou a bola para fora do terreno

$N_0 \ V \ N_1 \ Loc \ N_2$ sim

$N_0 \ V \ N_1$ sim

O João jogou fora o computador

$N_0 \ V \ N_1 \ Loc \ N_2$ não

$N_0 \ V \ N_1$ sim

Fixidez e liberdade

[O prefeito] põe [a foto] [no celular]

Análise distribucional:

*[Nhum] pôr [Nconc+Nhum] [Loc
(Nconc+Nhum)]*

Cada constituinte é **livre** (pode ter uma variedade de conteúdos)

O menino põe os óculos na mesa

O prefeito põe a mão no bolso

Sentido próprio: complementos livres

Sentido metafórico?

Fixidez e liberdade

[O prefeito] põe a mão no bolso

Sentido metafórico

Análise distribucional:

[O prefeito]? livre, Nhum

a mão? O prefeito põe os dedos no bolso sai da
metáfora: **fixo** com *pôr*

no bolso? O prefeito põe a mão na carteira sai
da metáfora: **fixo** com *pôr* e *mão*

Conclusão: *[Nhum] pôr a mão no bolso*

Fixidez e liberdade

[O prefeito] põe fogo nos documentos

[O prefeito]? livre, Nhum

*fogo? *O prefeito põe chama nos documentos.*

fixo com *pôr*

nos documentos? livre, em Nconc

Conclusão: *[Nhum] pôr fogo [em Nconc]*

Essa fórmula é superior à forma lexicográfica tradicional (*pôr fogo*), pois ela especifica o complemento livre

Fixidez e liberdade

[O João] engoliu um sapo

[O João]? livre, Nhum

*um sapo? O João engoliu uma rã sai da
metáfora: **fixo** com *sapo**

*um? O João engoliu (esse sapo + o sapo +
sapos + muitos sapos): livre, Det*

Conclusão: *[Nhum] engolir Det sapo*

Expressões fixas

Numa expressão fixa, os elementos geralmente não têm o sentido etimológico (*engolir um sapo*), mas podem ter (*pôr fogo*)

Uma expressão fixa pode ser mais frequente que um dos elementos (*batata baroa, jogar fora?*) ou menos frequente (*pôr a mão no bolso*)

Expressões fixas

Uma expressão fixa é geralmente mais metafórica do que os elementos (*engolir um sapo*), mas pode não ter diferença (*levar em conta*)

Bibliografia: Vale (2001); Baptista *et al.* (2004); Ranchhod (2005)

Propriedades sintático-semânticas

Informações sintático-semânticas sobre
entradas lexicais

Propriedades sintático-semânticas

$N_0 =: N_{hum}$

“O sujeito da construção de base pode denotar pessoas”

Jogar tem essa propriedade

Acontecer não tem

Valores de uma propriedade

Em geral, valor binário: “+” ou “-”

Às vezes, valor lexical: *Prep*

Propriedades sintático-semânticas

$N_0 =: N_{hum}$

$N_0 =: N_{conc}$

$N_0 =: N_{abs}$

$N_0 \text{ estar de } C \text{ Prep } N_1$

Prep

$N_1 =: N_{hum}$

$N_1 =: N_{conc}$

$N_1 =: N_{abs}$

Uma construção sintática codificada na forma de várias propriedades

Propriedades sintático-semânticas

Uma construção sintática é codificada na forma de várias propriedades:

- a fórmula
- a distribuição de cada elemento

Propriedades

- mais simples do que a construção
- reutilizáveis para outras construções, em outras tabelas...

Propriedades sintático-semânticas

Fórmula da propriedade $N_0 =: N_{hum}$
Documentação “O sujeito da
construção de base pode denotar pessoas”
‘The subject of the
basic construction can denote persons’

Fórmulas padronizadas por razões
computacionais

Propriedades sintático-semânticas

Fórmulas padronizadas

$N_0 =: N_{hum}$

“O sujeito da construção de base **pode assumir a forma de** um sintagma nominal denotando pessoas e respondendo à pergunta em *quem*”

O símbolo “=:” significa “pode assumir a forma de”

Construções de base e construções transformadas

Zé anuncia à imprensa que é candidato

= *Zé anuncia à imprensa ser candidato*

É a mesma entrada lexical *anunciar*

Uma das construções é descrita **em detalhes**

As outras **por diferença**

Construções de base e construções transformadas

O léxico-gramática explora a produtividade e a regularidade das transformações para **simplificar** a codificação

- 1) escolher um dos dois lados da fórmula (esquerdo ou direito) como “forma de base”
- 2) na descrição de cada item lexical, eliminar a descrição da outra forma (“forma transformada”)

Construções de base e construções transformadas

- 3) descrever como se prediz a forma transformada a partir da forma de base
- 4) na descrição de cada item lexical, especificar se a transformação se aplica ou não

Escolher a construção de base

A construção com o maior número de elementos

João está de acordo com Ana

João está de acordo

Escolher a construção de base

A construção mais frequente

João anuncia uma nova decisão

Uma nova decisão é anunciada por João

Pode não ser a mesma com todos os itens
lexicais, mas a decisão deve permanecer a
mesma

Um casal ocupava o carro

O carro estava ocupado por um casal

Propriedades sintático-semânticas

1) Construção de base

$N_0 =: N_{hum}$

$N_0 =: N_{conc}$

$N_0 =: N_{abs}$

N_0 estar de C Prep N_1

Prep

$N_1 =: N_{hum}$

$N_1 =: N_{conc}$

$N_1 =: N_{abs}$

2) Construções transformadas

N_0 estar de C

N_1 estar de C Prep N_0

Codificação das construções transformadas

Construção de base: N_0 estar de C Prep N_1

Construções transformadas

N_0 estar de C Prep $N_1 = \mathbf{N_0}$ **estar de C**

N_0 estar de C Prep $N_1 = \mathbf{N_1}$ **estar de C Prep N_0**

A fórmula da propriedade é a fórmula da
construção transformada

N_0 estar de C

N_1 estar de C Prep N_0

Tabelas de léxico-gramática

Registro das propriedades de cada entrada
lexical

Tabela de léxico-gramática

C	N0 =: Nhum	N0 =: Nconc	N0 =: Nabs	N0 estar de C Prep N1	Prep =: em	Prep =: com	Prep =: de	Prep =: por	Prep =: Loc	N1 =: Nhum	N1 =: Nconc	N1 =: Nabs	N1 =: V0-inf W	N1 estar de C Prep N0
acordo	+	-	+	+	-	+	-	-	-	+	-	+	-	+
caso	+	-	-	+	-	+	-	-	-	+	-	-	-	+
olho	+	-	-	+	+	-	-	-	-	+	+	+	-	-
quatro	+	-	-	+	-	-	-	+	-	+	+	-	-	-
recuperação	+	-	-	+	+	-	-	-	-	-	-	+	-	-
saco cheio	+	-	-	+	-	-	+	-	-	+	+	+	+	-
saco cheio	+	-	-	+	-	+	-	-	-	+	+	+	-	-
saída	+	-	-	+	-	-	-	-	+	+	+	-	-	-

Liliane A. Santana, 2011 “léxico” por causa dos itens lexicais
 “gramática” por causa das construções descritas

	A	B	C	D	E	F	G	H	I	J	K	
1	√ID^	<ENT>N	N0 =: Nhum	N0 =: N-hum	Det =: du	Det =: un-de Dnum Nmes	N0 être d'un certain N	N0 être de Det N Modif	N0 être de N Modif	il y avoir Det N Loc N0	N0 avoir le N de V0-inf W	Exemple Base1
110	109	cran	+	-	+	-	-	+	-	-	+	~
111	110	cubage	-	+	-	+	+	+	-	-	-	~
112	111	culture	-	+	-	-	-	-	+	+	-	Ce pays a une vieille culture
113	112	culture	+	-	+	-	+	+	+	-	-	~
114	113	dandysme	+	-	-	-	+	+	-	-	-	Max a un certain dandysme
115	114	datation	-	+	-	-	-	-	+	-	-	Ce vase a une datation
116	115	date	-	+	-	-	-	+	+	+	-	Cet événement a une date
117	116	débit	-	+	+	+	+	+	+	-	-	ce fleuve a un débit (*E+faible)
118	117	décharnement	+	-	-	-	+	+	-	-	-	Son visage a un certain décharner
119	118	décontraction	+	-	+	-	+	+	-	-	+	~
120	119	décor	-	+	-	-	-	+	-	+	-	(Ce palais+Cet opéra) a un magni
121	120	dégaine	+	-	-	-	-	+	-	-	-	~
122	121	délié	-	+	-	-	+	+	-	+	-	Ses gestes ont un délié ravissant
123	122	démarrage	-	+	-	-	-	+	+	-	-	Cette voiture a un démarrage diffic
124	123	démesure	-	+	+	-	+	+	-	+	-	Les décisions de Max ont une cer
125	124	démesure	+	-	+	-	+	+	-	-	+	Ce bâtiment a une certaine démes
126	125	densité	-	+	+	-	+	+	+	-	-	(Ce texte+Le brouillard) a une cert
127	126	dépravation	+	-	+	-	+	+	-	-	-	Max a une certaine dépravation

Tabela AN07 (Giry-Schneider, Balibar-Mrabti, 1993)

Classes

As entradas lexicais são classificadas em função das propriedades sintático-semânticas

A classificação utiliza as propriedades julgadas mais importantes:

- categoria gramatical
- número de argumentos
- preposições
- aceitabilidade de orações completivas nos argumentos...

Bibliografia: Leclère (2003)

Sumário

Objetivos históricos

Princípios

Modo de formalização

Resultados

Principais idiomas descritos

Francês

Coreano

Espanhol

Grego moderno

Italiano

Malgache

Português (sobretudo Portugal)

Predicados semânticos não verbais

A maioria dos predicados semânticos não são verbos, e sim:

- substantivos
- adjetivos
- expressões fixas

Independência dos argumentos

[O prefeito] anuncia [à imprensa] [que é candidato]

[Nhum] anunciar [a Nhum] [que F]

O conteúdo de cada argumento é independente do conteúdo dos outros

Ana anuncia a Zé que está grávida

Independência dos argumentos

O prefeito anuncia à imprensa que é candidato

O prefeito anuncia à imprensa que está grávida

O prefeito anuncia a Zé que é candidato

O prefeito anuncia a Zé que está grávida

Ana anuncia à imprensa que é candidato

Ana anuncia à imprensa que está grávida

Ana anuncia a Zé que é candidato

Ana anuncia a Zé que está grávida

Predicados semânticos não verbais

Em muitas frases, é melhor analisar como predicado semântico uma sequência que não se limita ao verbo

[O prefeito] põe fogo [nos documentos]

[João] tem saudade [do mar]

[João] está orgulhoso [do filho]

Construções com verbo suporte

João tem saudade do mar

Delimitação dos constituintes

*João tem do mar (*saudade + uma saudade imensa)*

Sei da saudade [que [João] tem [do mar]]

Esse é o mar [de que [João] tem [saudade]]

Conclusão: *[João] tem [saudade] [do mar]*

Construções com verbo suporte

[João] tem [saudades] [do mar]

Análise distribucional

*[Nhum] ter [Nabs] [de (N + ?que F + ?que
Fsubj + quando F)]*

Independência dos constituintes?

João tem saudades (do mar + do nariz)

João tem vergonha (?do mar + do nariz)

*João tem infecção (*do mar + do nariz)*

Existe interdependência entre o conteúdo de dois constituintes

Construções com verbo suporte

Com *saudade*:

[Nhum] ter [saudade] [de (N + ?que F + ?que Fsubj + quando F)]

Com *infecção*:

[Nhum] ter [infecção] [(de + Loc) Npc]

A sequência *ter saudade* ou *ter infecção* cumpre uma função comparável à de um verbo como *jogar* ou *falar*

Considerar o substantivo como fixo para fazer a análise distribucional

Construções com verbo suporte

[João] tem saudade [do mar]

[Nhum] ter saudade [de (N + que F + que Fsubj)]

[João] tem infecção [na garganta]

[Nhum] ter infecção [(de + Loc) Npc]

[João] tem gripe

[Nhum] ter gripe

[João] tem inveja [de Zé]

[Nhum] ter inveja [de (N + que F)]

Construções com verbo suporte

[João] tem infecção [na garganta]
[Nhum] ter infecção [(de + Loc) Npc]

ter: verbo suporte (*ter, dar, estar, ser, tomar, fazer, bater, estar com, estar em...*)
infecção: substantivo predicativo (*Npred*)

[João] dá uma limpeza [no chão]
[Nhum] dar Det limpeza [Loc Nconc]

Construções com verbo suporte

[João] é fiel [aos princípios]

[Nhum] ser fiel [a (Nhum + Nabs)]

[João] está orgulhoso [do filho]

[Nhum] estar orgulhoso [de (N + que F + que Fsubj) + com N]

[Essa bebida] é estranha

[N + que F + que Fsubj] (ser + estar) estranho

Bibliografia: Harris, 1964; Gross, 1981

Um critério sintático

Uma construção com verbo suporte (CVS)
sempre tem uma variante na qual o verbo
suporte é ausente, sem modificação
imprevisível do sentido

João deu um carinho no cachorro

*Depois de (E + dar) um carinho no cachorro,
João começou a espirrar*

O carinho de João no cachorro o acalmou

Redução de verbo suporte

É importante **conhecer** as variantes de CVS na qual o verbo suporte é ausente, para poder **reconstituir** em novos exemplos e **testar**

O carinho de João no cachorro o acalmou

Redução de verbo suporte

João deu um carinho no cachorro. Isso o acalmou

[Formação de oração relativa]

O carinho que João deu no cachorro o acalmou

[Redução de verbo suporte]

O carinho de João no cachorro o acalmou

Distribuição do verbo suporte

Depende do predicado semântico

João tem sorte

* *João há sorte*

João perdeu a sorte

Tem árvores na praça

Há árvores na praça

Nunca tinha pensado

Nunca havia pensado

Expressões fixas

As expressões fixas são mais numerosas do que as palavras simples

O léxico-gramática do francês registra

13 900 entradas lexicais de verbos

39 600 de expressões fixas

Irregularidades

O léxico é caótico

As diferenças entre entradas lexicais são subestimadas

As propriedades dependem da entrada lexical

Sua besta quadrada!

?* *Sua estrela!*

Irregularidades

No caso de uma palavra ambígua, uma entrada lexical a cada sentido:

João pulou na piscina

João deu um pulo na piscina

João pulou uma pergunta

* *João deu um pulo numa pergunta*

Morfologia e sintaxe

As derivações produtivas são transformações sintáticas

brasileiro/abrasileirar, francês/afrancesar, português/aportuguesar...

Produtividade:

- a quantidade de entradas lexicais afetadas
- a reprodução das mesmas propriedades com cada entrada (regularidade)

A produtividade pode não ser total

*inglês *ainglesar anglicisar*

Morfologia e sintaxe

É sempre melhor expressar as formas como sendo frases

Ana é brasileira/O fato de morar ali a brasileira

Ana

Ana é francesa/O fato de morar ali a francesa

Ana

Ana é portuguesa/O fato de morar ali

a portuguesa Ana

As frases têm um sentido mais preciso

Morfologia e sintaxe

N₀ ser Adj/N_{caus} a-Adj-ar N₀

Uma transformação sintática no sentido de
Z. Harris

Reprodução dos mesmos fatos com cada
adjetivo étnico:

As duas frases não têm o mesmo sentido

A **diferença de sentido é a mesma** com
cada adjetivo étnico

Morfologia e sintaxe

N_0 *ser* *Adj*/ N_{caus} *a-Adj-ar* N_0

Possibilidade de “aplicar” a relação:

Todos os elementos da fórmula são

- **variáveis** (N_0 , *Adj*, N_{caus}) que podem ser preenchidas por um material lexical variado, mas pelo mesmo material à esquerda e à direita
- ou **constantes** gramaticais e morfológicas (*ser*, *a-*, *-ar*)

Bibliografia: Harris, 1964; Gross, 1981

Conclusão

Objetivo de cientificidade

Métodos oriundos das ciências experimentais

Resultados teóricos

Aplicações práticas: dicionários e gramáticas
para processamento das línguas

Bibliografia

- Baptista, Jorge; Correia, Anabela; and Fernandes, Graça. 2004. "Frozen Sentences of Portuguese: Formal Descriptions for NLP". In T. Tanaka, A. Villavicencio, F. Bond, A. Korhonen (eds.), *2nd ACL Workshop on Multiword Expressions: Integrating Processing (MWE'04)*, pp. 72-79, <http://acl.ldc.upenn.edu/acl2004/mwe/pdf/baptista.pdf>
- Bloomfield, Leonard. 1933. *Language*, <http://books.google.com.br/books?id=Gfrd-On5iFwC>

Bibliografia

Boons, Jean-Paul; Guillet, Alain; Leclère, Christian. 1976. *La structure des phrases simples en français: constructions intransitives*. Genève: Droz, Introdução, pp. 29-52, <http://books.google.fr/books?id=BlNZtt2H6QoC>

Giry-Schneider, Jacqueline; Balibar-Mrabti, Antoinette. 1993. *Classes de noms construits avec avoir*. Rapport technique 42, LADL, Université Paris 7.

Bibliografia

Gross, Maurice. 1975. *Méthodes en syntaxe. Régime des constructions complétives*, Paris: Hermann.

Gross, Maurice. 1981. Les bases empiriques de la notion de prédicat sémantique, *Langages* 63, Paris: Larousse, pp. 7-52,
http://www.persee.fr/articleAsPDF/lgge_0458-726x_1981_num_15_63_1875/article_lgge_0458-726x_1981_num_15_63_1875.pdf

Bibliografia

Guillet, Alain. 1986. "Représentation des distributions dans un lexique-grammaire". *Langue Française* 69, Paris: Larousse, pp. 85-107,
http://www.persee.fr/web/revues/home/prescript/article/lfr_0023-8368_1986_num_69_1_6364

Bibliografia

Harris, Zellig. 1964. *The Elementary Transformations*. Transformations and Discourse Analysis Papers, no. 54, Philadelphia: University of Pennsylvania, 117 pages (reimpr. em 1970: *Papers in Structural and Transformational Linguistics*. Dordrecht/Holland: D. Reidel, pages 482-532).

Bibliografia

Leclère, Christian. 2003. "The lexicon-grammar of French verbs: A Syntactic Database". in *Proceedings of the First International Conference on Linguistic Informatics*, Kawaguchi Y. et alii (eds.), UBLI, Tokyo University of Foreign Studies, pp. 33-46, <http://hal.archives-ouvertes.fr/hal-00192888/en/>

Bibliografia

- Montague, Richard. 1970. "Universal grammar".
Theoria 36:373–398 (reimpr. en 1974 :
Thomason, Richmond, ed., *Formal
Philosophy. Selected Papers by Richard
Montague*. New Haven).
- Popper, Karl. 1959. *The Logic of Scientific
Discovery*. London: Routledge,
<http://books.google.com.br/books?id=Yq6xeupNStMC>

Bibliografia

Ranchhod, Elisabete M. 2005. "Using Corpora to Increase Portuguese MWU Dictionaries. Tagging MWU in a Portuguese Corpus". In P. Danielsson, M. Wagenmakers (eds.), *Corpus Linguistics* (CL'05), 14 pp., <http://infolingu.univ-mlv.fr/english/Bibliographie/Ranchhod-mwu-corpuslinguistics05.PDF>

Bibliografia

Tesnière, Lucien. 1959. *Éléments de syntaxe structurale*. Paris : Klincksieck.

Vale, Oto Araújo. 2001. "Transparência e opacidade de expressões cristalizadas". In Hirata-Vale, Flávia B.M. (eds.), *Anais do IV Seminário Nacional de Literatura e Crítica e do II Seminário Nacional de Lingüística e Língua Portuguesa*, Goiânia Gráfica e Editora Vieira, pp. 240-246.