

Linguagem

Origens, Evolução e Aprendizado



por Sergio Navega, Publicações Digitais Intelliwise

Este documento foi originalmente publicado como material didático de um dos seminários que compõe o "Curso de Formação Crítica de Analistas do Conhecimento". Esta versão foi especialmente formatada para ser distribuída através das **Publicações Digitais Intellwise**. Alguns dos tópicos desenvolvidos neste material podem ficar um pouco obscuros, já que foram originalmente concebidos como suporte para uma exposição oral de mais de 3 horas. Se tiver qualquer dúvida, durante o Período de Acesso deste material você poderá fazer quantas perguntas quiser ao instrutor. Para imprimir este material, use papel em formato carta.

© 2000, 2003 Sergio Navega, Intellwise Research
Edição original de Setembro de 2000
1a. Edição Digital em Fevereiro de 2003

Nenhuma parte deste material pode ser copiada, armazenada ou transmitida por quaisquer que sejam os meios sem prévia autorização por escrito da Intellwise Research and Training. Você pode imprimir este material para uso pessoal. Se quiser utilizar partes deste material para finalidades didáticas, consulte-nos para obter permissão. O autor esforçou-se para produzir um material livre de erros e defeitos. No entanto, nem o autor nem a Intellwise Research poderão ser responsabilizados pelo uso que se fizer das informações aqui contidas.

Em caso de dúvidas sobre alguns dos pontos deste material, você pode contatar Sergio Navega através do E-mail: snavega@attglobal.net. Não se esqueça de citar o título do slide sobre o qual você quer explicações. A home page de Sergio é <http://www.intellwise.com/snavega>

Para maiores informações sobre outras
Publicações Digitais Intellwise veja esta página
<http://www.intellwise.com/digital>

Este seminário pode ser proferido para um grupo fechado
em sua empresa. Acesse a página abaixo para mais informações
<http://www.intellwise.com/seminars>

Linguagem - Origens, Evolução e Aprendizado

CONTEÚDO

- ☑ Elementos das Linguagens
 - Os Passos de Uma Comunicação
 - Comunicações Não-Simbólicas
 - Referentes, Significados, Sentidos

- ☑ Semântica
 - Reverendo Noções Ultrapassadas

- ☑ Estruturas da Linguagem
 - Hierarquia, Sintaxe, etc
 - Gramáticas Descritivas e Prescritivas

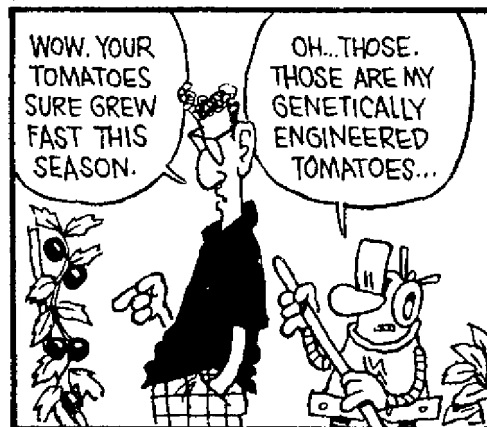
- ☑ Geratividade da Linguagem

- ☑ Psicolinguística e Revolução Chomskyana

- ☑ Linguagem: Instinto Humano?
 - Linguagem de Signos Nicaraguense

- ☑ Repensando o Que é Inato

Linguagem É Essencialmente Humana



ROBOTMAN® by Jim Meddick

Não há dúvidas de que Linguagem é uma das principais características daquilo que nos torna humanos


Linguagem: Para Que Serve?

A comunicação entre agentes informacionais

Ser Humano - Ser Humano

Ser Humano - Máquina

Máquina - Máquina

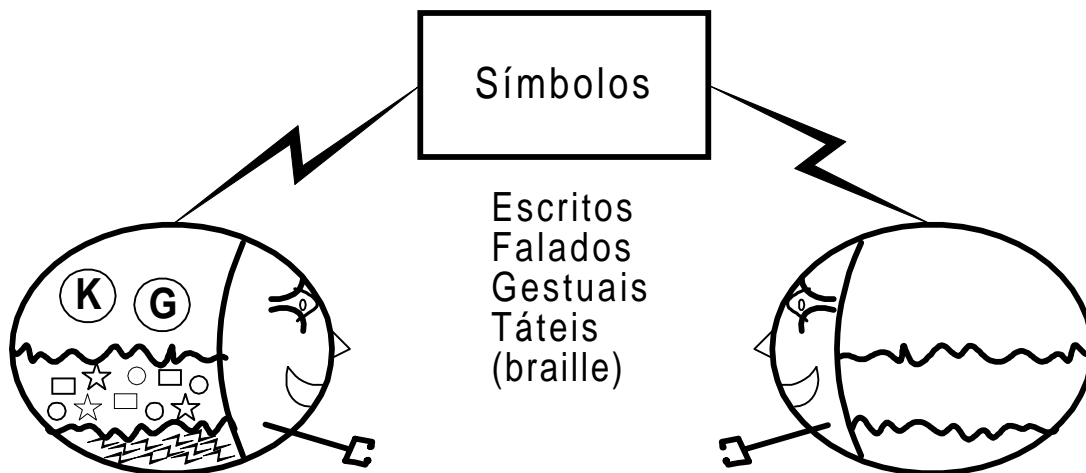
 Ser Humano Consigo Mesmo

► Linguagens Naturais

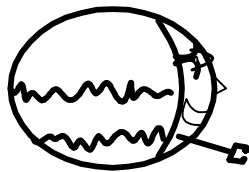
Português, Inglês, Francês, etc

► Linguagens Formais (Artificiais)

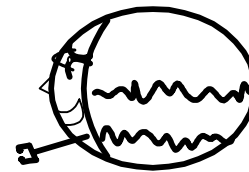
LISP, Lógica de Primeira Ordem (FOPC), etc



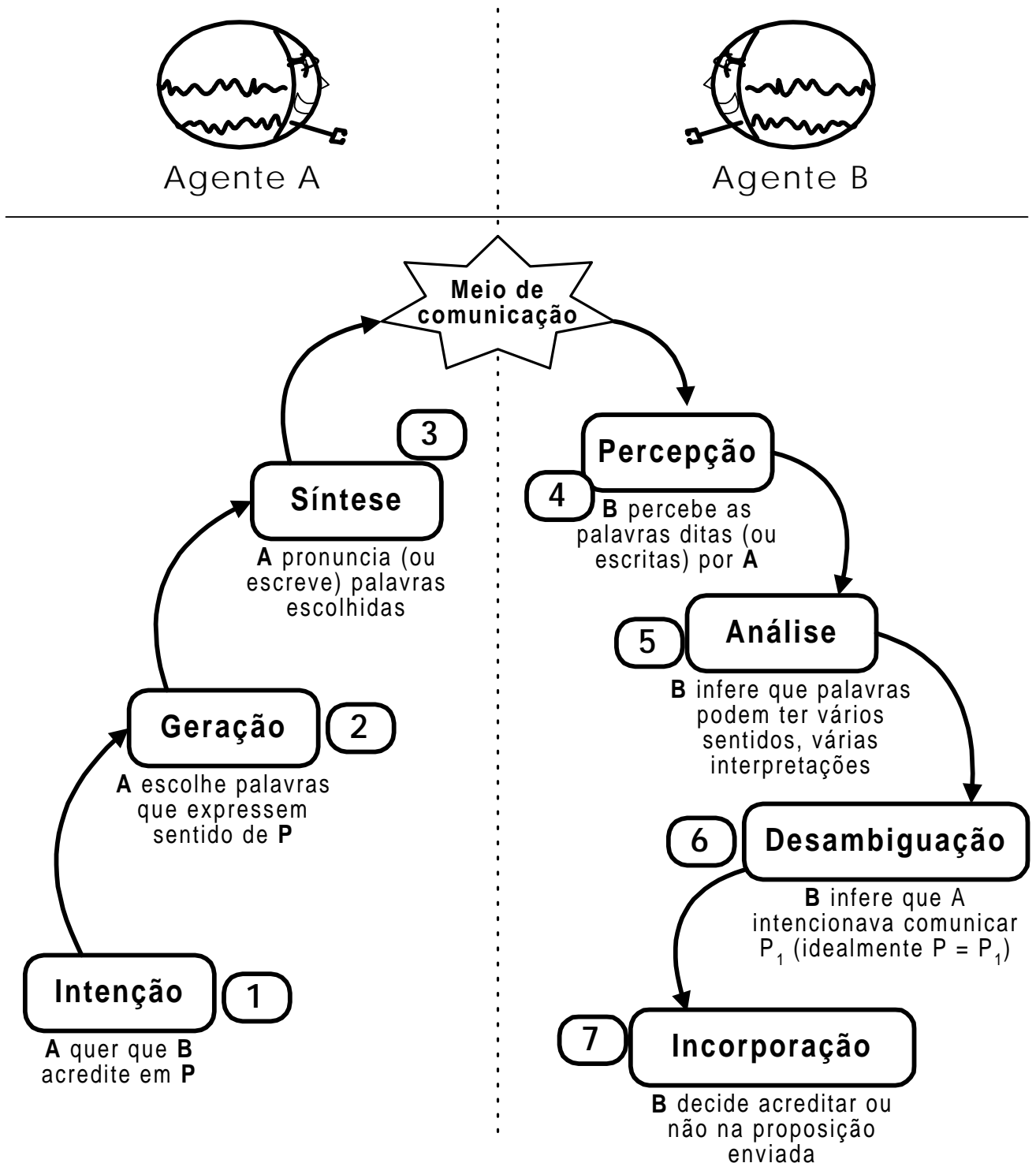
Os Passos de Uma Comunicação



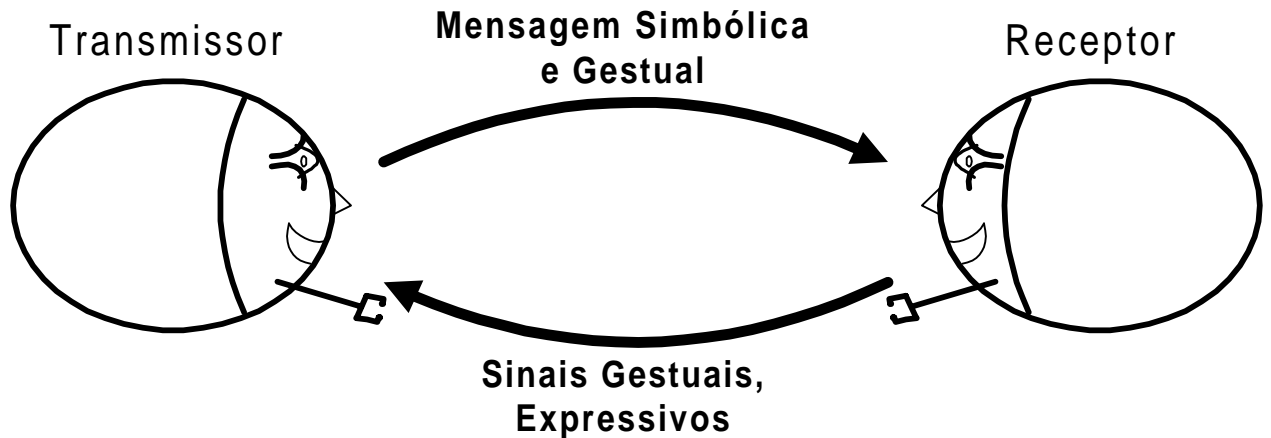
Agente A



Agente B



Comunicação Não-Simbólica



▶ Aspectos Não Simbólicos Usados Durante Comunicação, a Linguagem Corporal

▶ Gestos, sinais, expressões, murmúrios

▶ Tipos de mensagens transmitidas

- Eu já sei disso
- Não estou certo, convença-me
- Não estou interessado nisso
- Muito Interessante!
- Mais devagar
- Mais depressa
- E daí?
- Não concordo
- Que estranho. Tem certeza?
- Tenho algo a acrescentar sobre isto
- Entendi (hum, hum...)

▶ Essa sinalização melhora a relação sinal/ruído além de otimizar a transmissão de informação

Referentes, Significados e Sentidos

Trator



Flor

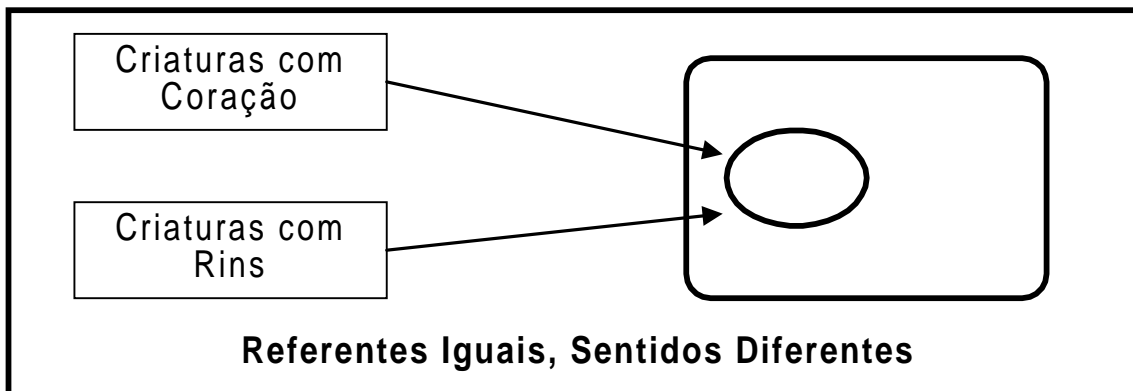
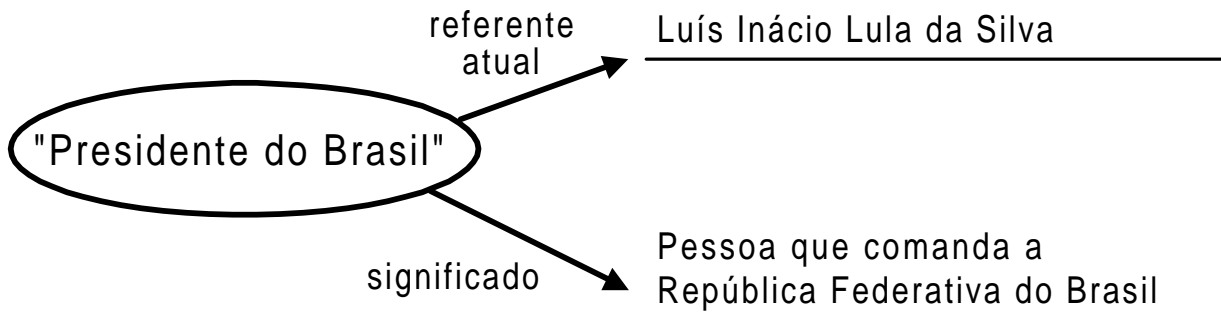


Leitura

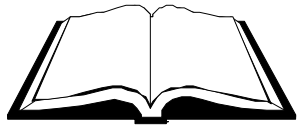


"atividade sendo executada durante a interpretação desta sentença"

Referentes



Palavras, Conceitos, Semântica

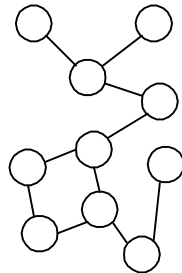


L I V R O

"sensação do
gosto do mel"



? ? ? ?

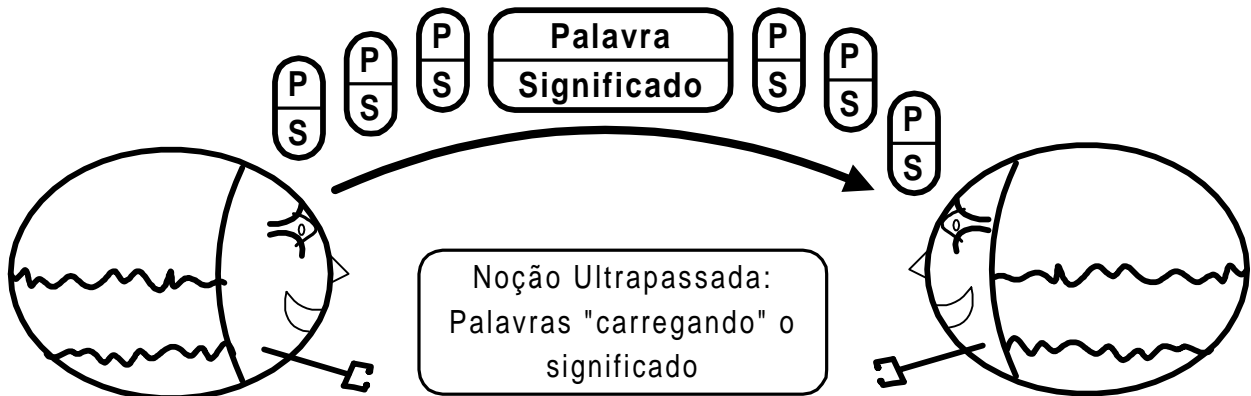


J U S T I Ç A

O porquê dos problemas da IA simbólica

- ☹️ Modelagem do conhecimento a partir dos símbolos linguísticos, e não das essências semânticas dos conceitos (o aprendizado dos bebês)
- ☹️ Desconsidera os mecanismos de desenvolvimento e estruturação dessas noções semânticas (similaridade, repetição, cópia analógica, etc), necessários sempre que se precisa adquirir conhecimento novo

Noções Ultrapassadas de Linguagem



Por Que Não É Assim?

O significado (semântica) das palavras **tem que estar no receptor**, para que possa ser compreendido. Frases tem sentido diferente de acordo com quem as ouve (ou seja, de acordo com a construção semântica individual que cada pessoa tem).

Mesmas Palavras, Sentido Diferente

Os políticos pensam que o público não sabe

O público sabe que os políticos não pensam

As Múltiplas Estruturas da Linguagem

Pragmática

Estudo dos fatores que governam nossas opções de usos da linguagem em contextos diversos. Regras sociais de uso da linguagem.

Semântica

Estudo do significado dentro da comunicação linguística

Sintaxe

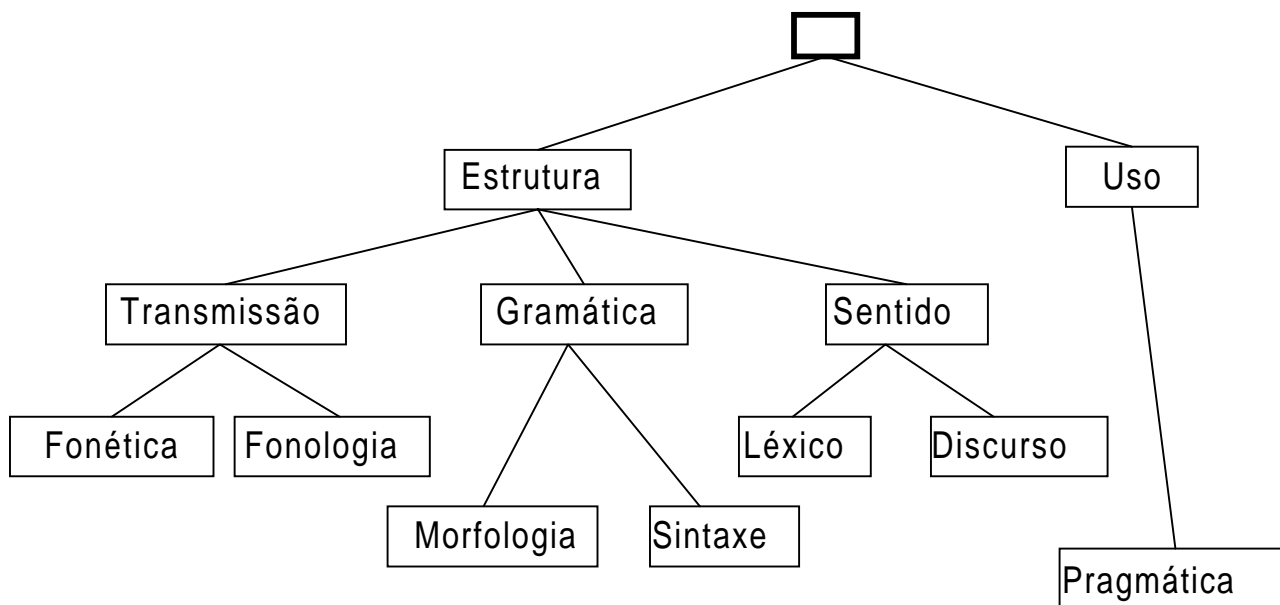
Estudo das relações entre palavras, os princípios e processos através dos quais sentenças são formadas

Morfologia

Estudo da estrutura, formação e classificação das palavras

Fonética (Fonologia)

Estudo de como os sons são produzidos, transmitidos, recebidos e percebidos



Divisões de Acordo Com a Semiótica



Sintaxe

Relação dos signos com outros signos, formando expressões linguísticas



Semântica

Relação dos signos com os objetos do mundo exterior ao qual se referem



Pragmática

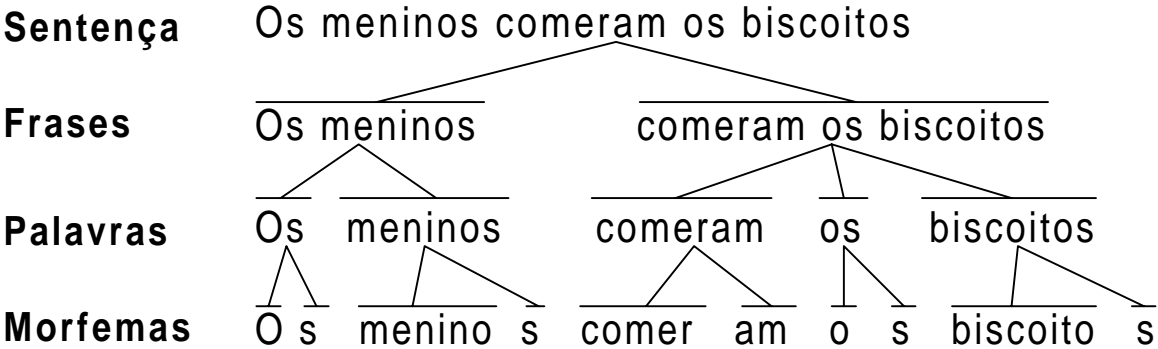
Relação entre signos e suas expressões aos usuários (intérpretes), incluindo relações de ênfase e situação social em que são usadas

Uma visão diferente de semântica e sua importância na memória humana

qui caqui quiu
quico que
quico que caque
caqui

Como decorar a sequência acima? Lembrar-se de um sentido
"que caqui que o Kiko quer? Kiko quer qualquer caqui"

Linguagem é Hierárquica



Outro Comentário Sobre Semântica

Entendendo Semântica como "execução" de um programa por processadores diferentes

- Pentium
- Motorola MC68000
- Zilog Z80
- IBM S390

A = B + C

Sintaxe

```
LD    AX, A
LD    BX, B
ADD   AX, BX
LD    C, AX
```

Semântica

As Regras de Sintaxe

Com 20 palavras podemos formar 100.000 trilhões de sentenças diferentes (10^{20}). Apenas um minúsculo subset é gramatical

O menino chutou a bola

* O menino bola chutou a

Um passarinho entrou em casa

* Um passarinho casa em

* Um casa entrou passarinho em

Sintaxe: O conjunto de regras das formas permitidas. Sentenças que obedecem a essas regras são ditas gramaticais

Não confundir gramática com aceitabilidade em termos de sentido

Colorless green ideas sleep furiously (Chomsky)

Idéias verdes incolores dormem furiosamente

*Furiosamente incolores idéias dormem verdes

Uma Gramática Simples

Sentença -> FN + FV

FN -> Art + Subs

FV -> Verbo + FN

Art -> o, a, os, as

Subs -> homem, bola, cadeira, mesa

Verbo -> chutar, pegar, fazer, sentar

Sentença

FN + FV

Art + Subs + FV

Art + Subs + Verbo + FN

O + Subs + Verbo + FN

O + homem + Verbo + FN

O + homem + chutou + FN

O + homem + chutou + Art + Subs

O + homem + chutou + a + Subs

O + homem + chutou + a + bola

Sujeito, Verbo, Objeto

SVO

O menino viu o homem

OVS

João convidei, Joaquim eu não

VSO

Governe você seu reino

SOV

Pensantes poetas dolorosas virgens mantém



Alguns exemplos de
OSV pelo mestre Yoda

Doente eu fiquei

Com a Força eu estou forte

Seu pai ele é

Quando 900 anos de idade
você atingir, parecer tão
bem você não vai

Gramáticas Descritivas e Prescritivas

Gramáticas Descritivas

São um registro de como as pessoas costumam falar e escrever; descrevem como o sistema funciona na prática

"Who did you sit with?"

"Eu não vi nada" (duplo negativo)

Gramáticas Prescritivas

São especificações de como se deveria usar a linguagem, em geral seguindo normas estabelecidas por autoridades no assunto

"With whom did you sit?"

"Eu nada vi" ou "Eu não vi"

"This is a rule up with which we should not put"
Winston Churchill

"This is a rule which we should not put up with"

Linguagem é Gerativa

antinconstitucionalissimamente
anti in constitu[i] cional issima mente

carta	cartas
casa	casas
*cerprito	cerpritos
avião	aviões
chavão	chavões
*gotibão	gotibões

Geratividade das linguagens naturais acontecem porque temos tendência a formar regras genéricas, que podem ser empregadas tanto para gerar como para reconhecer

colheres de metal são "nãomicroondáveis"
Interessante o fato de que crianças são as mais criativas no uso do poder gerativo das regras linguísticas

Recursividade (e os limites da mente humana)

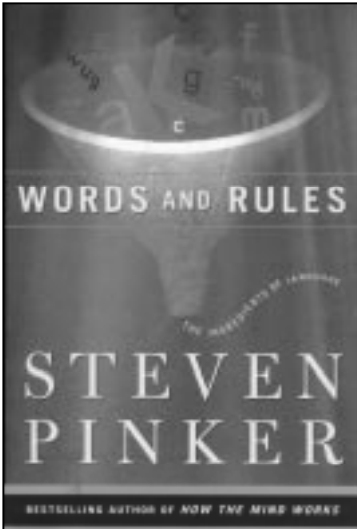
- ▶ Ela é competente
- ▶ Eu acredito que ela é competente
- ▶ João falou que eu acredito que ela é competente

- ▶ O diretor não aceitou a tese de Mário de que João disse que eu acredito que ela é competente para o cargo postulado por Rodrigo

O Caso dos Verbos Regulares em Inglês

Um dos fenômenos mais estudados pelos cientistas cognitivos

A	{	climb	climbed	climbed
		play	played	played
		ring	ringed	ringed
		blink	blinked	blinked
B	{	go	went	gone
		run	ran	ran
C	{	bring	brought	brought
		think	thought	thought
		sink	sank	sunk

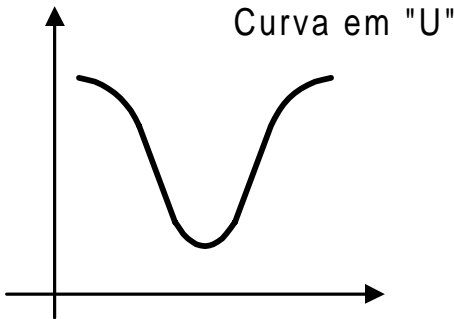
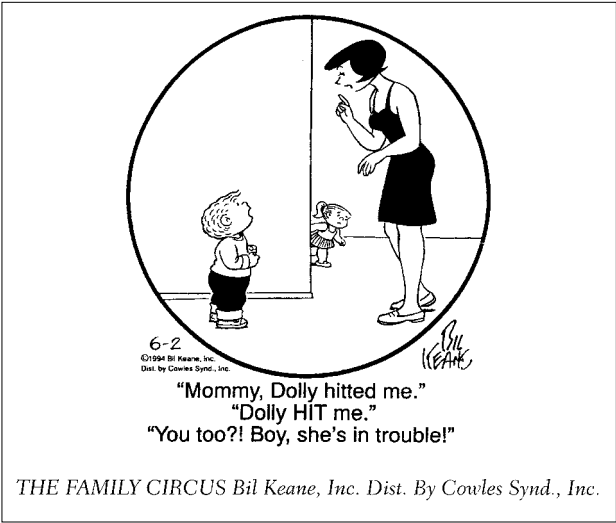


Overregularization

* "Yesterday, we goed"

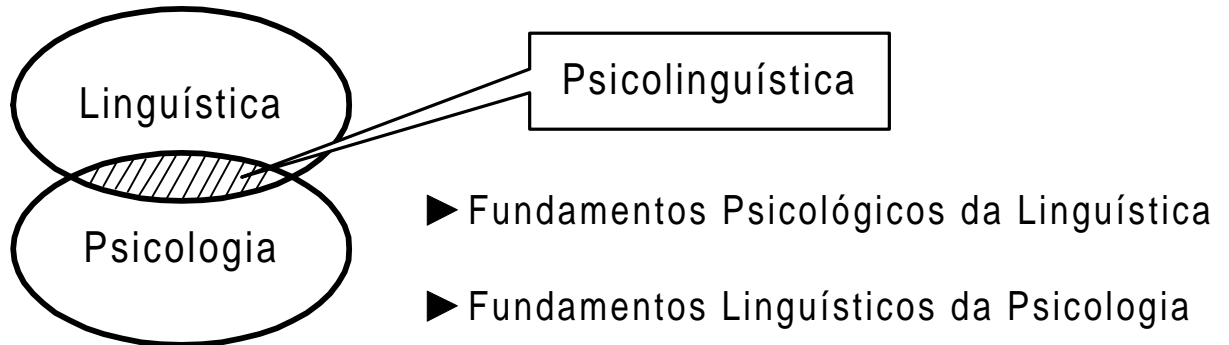
Falado alguns meses após ter usado a forma correta
"Yesterday, we went"

Não há esquecimento das formas irregulares, mas sim uma "briga" entre estas e as regras das formas regulares



Períodos de confusão onde a criança combina essas táticas e gera coisas como:
"He wented"
(Marcus, Pinker et al. 1992)

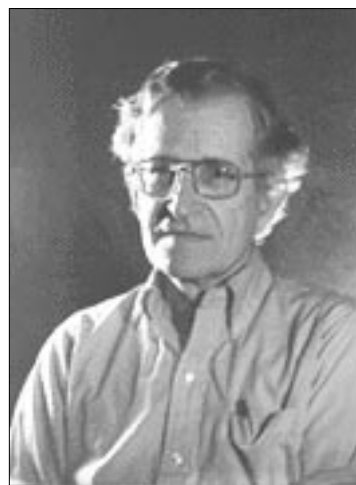
O Que É Psicolinguística



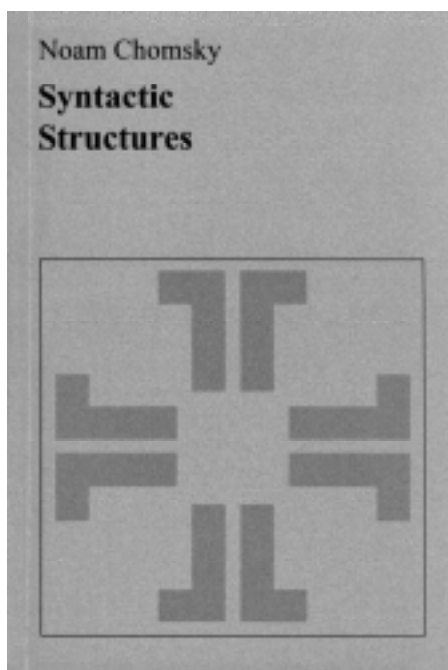
- 👉 Começou preocupando-se essencialmente com as palavras (memorização de listas, lembrança de pares de palavras)
- 👉 Nessa época, o tamanho do vocabulário de uma pessoa era tido como indicador de sua idade mental. Usava-se testes de vocabulário para determinar a inteligência das pessoas
- 👉 Com o tempo, os psicolinguistas foram verificando que a linguagem humana era por demais complexa para ser explicada pelas teorias de aprendizado baseadas no condicionamento behaviorista (Skinner). O ambiente estava propício para a revolução que viria em seguida

A Revolução Na Linguística

- Chomsky percebeu que era possível separar a estrutura das sentenças (gramática) de seu significado (semântica) e da forma de usar (pragmática). Propôs estudar essa estrutura como fenômeno isolado, adicionando um tratamento matemático e lógico, transformando a linguística em uma ciência da sintaxe
- Confrontou o behaviorismo (Skinner)
- Gramáticas Universais e Gerativas
- O "problema da pobreza dos estímulos"
- A Hipótese do "órgão de linguagem", de origem genética (inata)



Avram
**Noam
Chomsky**
(1928 -)



Mudou o foco da Psicolinguística de palavras (vocabulário) para a gramática

Saussure, Leonard Bloomfield, Zellig Harris
A Tese de Chomsky (Transformational Analysis), indo contra Harris, seu orientador

Syntactic Structures foi escrito em 1957, com o próprio Chomsky custeando parte de sua publicação. Deste essa época, nunca esteve fora de circulação (minha cópia é a 12ª edição)

Outros Caminhos da Linguística Moderna

As idéias de Chomsky, embora tenham tido influência fundamental em praticamente todas as áreas da linguística moderna, receberam diversas críticas e complementações. Uma das críticas comuns decorre da fixação em gramática pelos chomskianos, com conseqüente desatenção pela semântica

Regra R1

Sentenças tem uma frase nominativa como sujeito seguidas por uma frase verbal como predicado

A mulher sorriu

*A mesa sorriu

Para Chomsky, importam apenas as relações gramaticais (sintáticas).

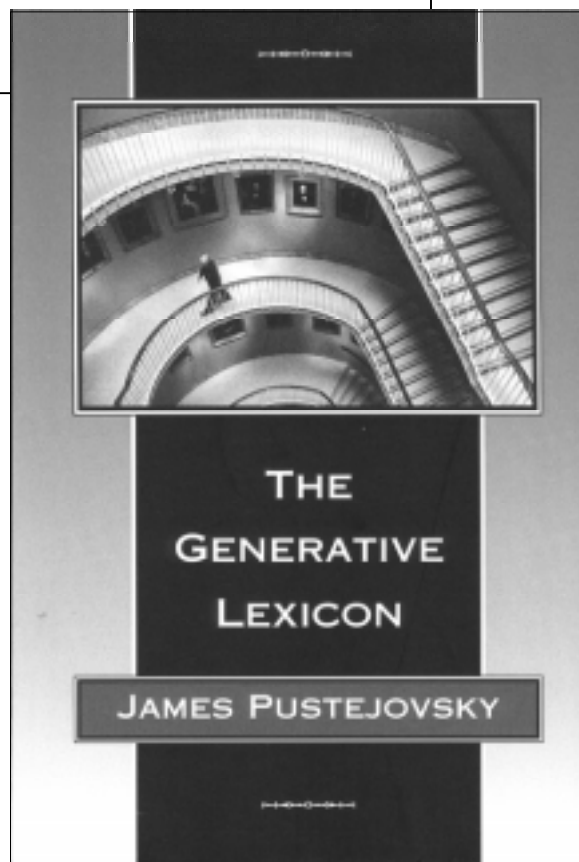
Ambas as frases são gramaticalmente corretas, mas a segunda não faz sentido

Uma das formas de introduzir a semântica é através de Léxicos

Léxico

Sorrir { Requer um ser animado,
possivelmente humano

Formas de processamento de linguagem natural através de métodos estatísticos



James Pustejovsky e os Léxicos Gerativos

Linguagem: Instinto Humano?

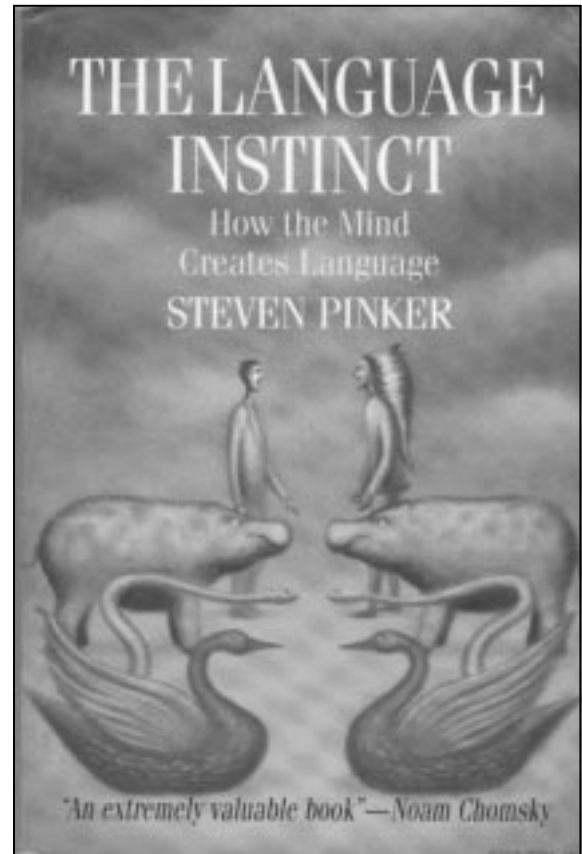
- Específica da espécie
- Uniformemente distribuída
- Criação espontânea
- Habilidade destacada de outras
- Período de aprendizagem específico

Síndrome de Down

"Maria está chora na igreja"

"Os rapazes come quatro bolo"

SLI - Specific Language Impairment



Síndrome de Williams

Raro defeito genético (1 em cada 20.000 nascimentos)

- Baixo QI
- Habilidades matemáticas ruins
- Má localização espacial (raciocínio visual)
- Boa linguagem, frequentemente bons em música

O Nascimento de Uma Linguagem

1980: Crianças surdas de Nicarágua desenvolvem por si próprias uma linguagem baseada em sinais, sem contato ou influência de nenhuma das outras 200 linguagens de sinais existentes no mundo



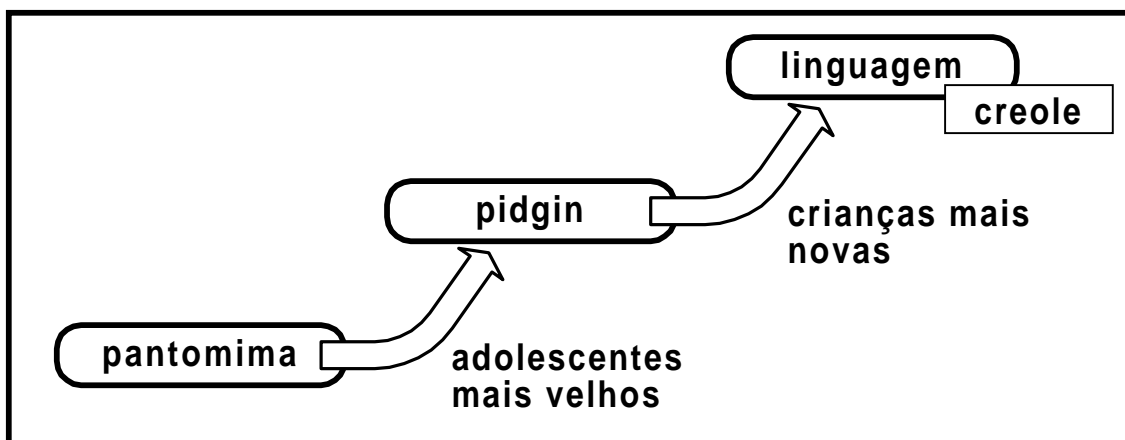
Yuri Mejia (9 anos) contando a estória do elefante Babar, andando nas costas de sua mãe que é atingida por um tiro. Babar foge e depois se entrega às lágrimas



Centenas de crianças surdas, provenientes de diversas partes da Nicarágua, foram trazidas a Manágua para estudarem em 2 escolas especiais. Cada criança dispunha de um "estoque" de sinais simples (mímicas) diferentes entre si. Mas ao serem colocadas em conjunto, alguns desses sinais se estabeleceram entre elas. Seus professores, espantados, ficaram acompanhando o nascimento daquilo que se tornaria uma nova linguagem: ISN (Idioma de Signos Nicaraguenses)

O Fenômeno Nicaraguense

- ▶ ISN possui mesmos princípios universais, conforme preconizados por Chomsky
- ▶ Steven Pinker considera esse evento como uma prova da existência de módulos "hard-wired" no cérebro das pessoas para produção/entendimento de linguagem (há cientistas que não pensam assim)
- ▶ Estudantes mais velhos (adolescentes) já apresentavam diferenças em relação aos mais novos (5 a 6 anos). Os mais novos tinham sinais com maior número de variantes, adicionando sofisticação aos sentidos possíveis
- ▶ Os mais novos aprenderam rapidamente a linguagem dos mais velhos e inconscientemente enriqueceram a língua



Esperanto (criado por adultos) é menos sofisticado do que o ISN

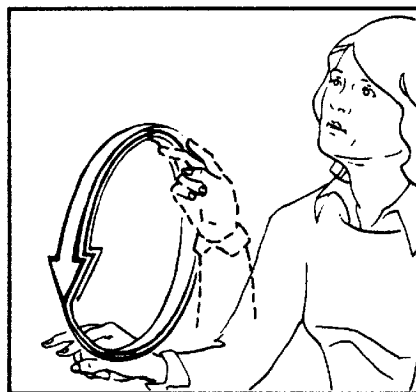
"Nenhum linguista pode criar uma linguagem com metade da complexidade e riqueza existente naquela concebida por uma criança de 4 anos de idade"

Judy Kegl, Northeastern University

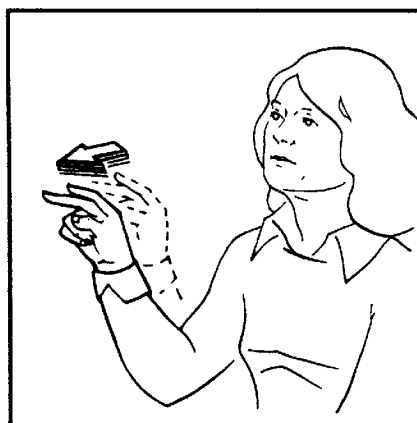
Modificação de Características de Um Verbo



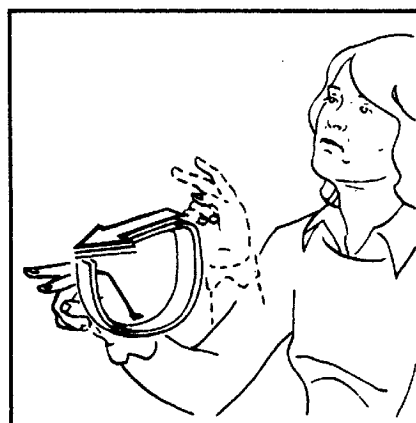
'Look at'



'Look at for a long time'



'Look at incessantly'



'Look at repeatedly'

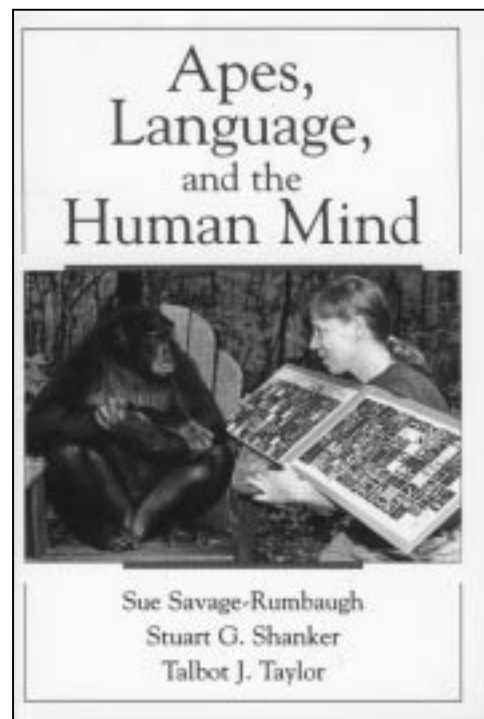
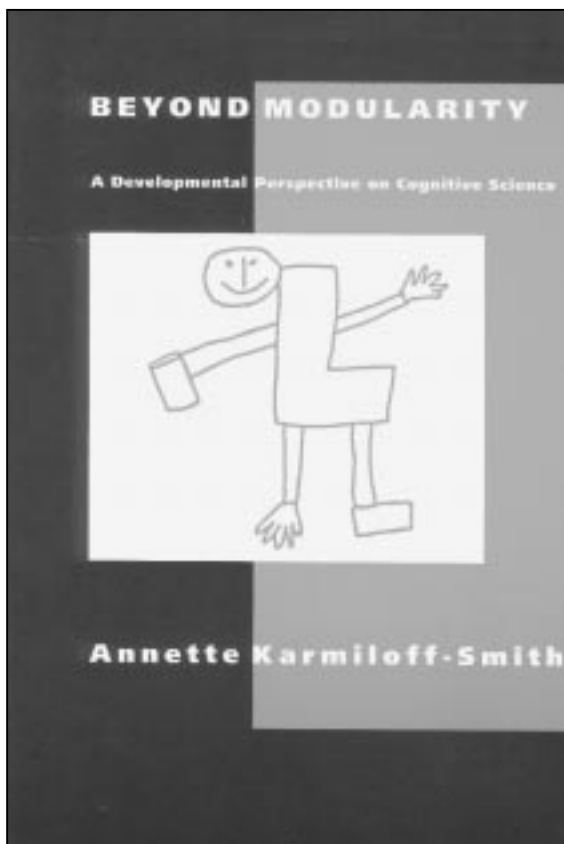
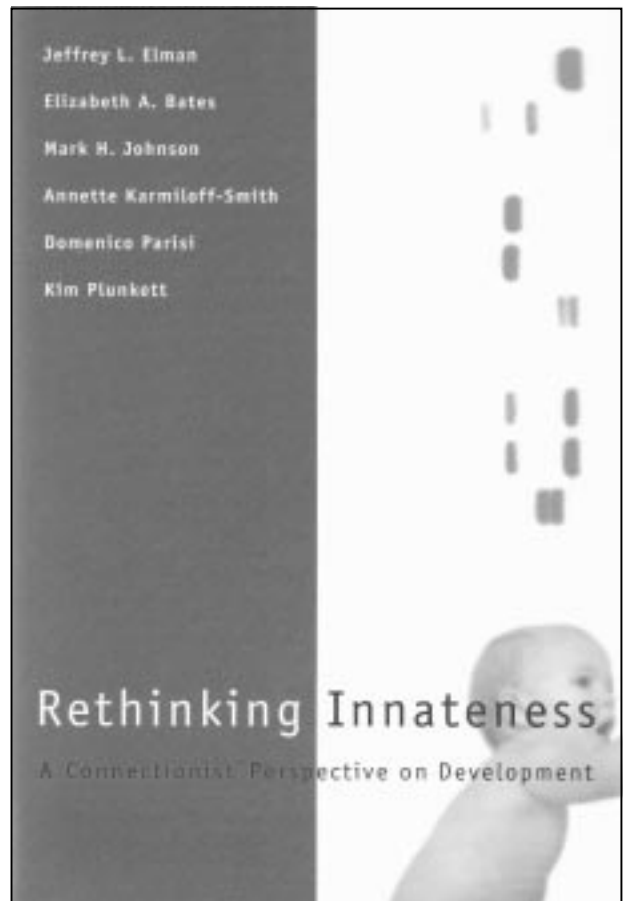


'Look at habitually'

Jackendoff (1994)

Repensando o Que É Inato

- Específico da espécie?
Kanzi e outros Bonobos
- Pobreza de estímulos?
Está mais para excesso deles
- Deficiências genéticas?
Todas afetam também outras modalidades (KE family)
- Específicas ao domínio?
Não há evidências conclusivas
- Modelos não inatos?
Há diversos modelos conexionistas interessantes



Referências

- Altman, Gerry T. (1997) *The Ascent of Babel*. Oxford University Press.
- Charniak, Eugene (1993) *Statistical Language Learning*. Bradford Book, MIT Press
- Chomsky, Noam (1957) *Syntactic Structures*. Mouton, The Hague-Paris (12th printing, 1976)
- Cobley, Paul; Jansz, Litza (1999) *Introducing Semiotics*. Totem Books.
- Crystal, David (1996) *The Cambridge Encyclopedia of the English Language*. Cambridge University Press.
- Crystal, David (1997) *The Cambridge Encyclopedia of Language*. Cambridge University Press.
- Elman, Jeffrey; Bates, Elisabeth; Johnson, Mark H.; Karmiloff-Smith, Annette; Parisi, Domenico;
Plunkett, Kim (1996) *Rethinking Innateness, A Connectionist Perspective on Development*. Bradford Book, MIT Press.
- Jackendoff, Ray (1994) *Patterns in the Mind*. BasicBooks.
- Maher, John; Groves, Judy (1997) *Introducing Chomsky*. Totem Books.
- Miller, George A. (1996) *The Science of Words*. Scientific American Library.
- Osborne, Lawrence (1999) *A Linguistic Big Bang*. New York Times Magazine
- Pinker, Steven (1994) *The Language Instinct*. William Morrow and Company, Inc
- Pinker, Steven (1999) *Words and Rules*. Basic Books.
- Pustejovsky, James (1995) *The Generative Lexicon*. MIT Press.
- Reisberg, Daniel (1997) *Cognition, Exploring the Science of Mind*. W. W. Norton & Company.
- Russell, Stuart; Norvig, Peter (1995) *Artificial Intelligence, A Modern Approach*. Prentice-Hall Inc