Hands on / Mãos na massa:
fundamentos e dicas de uma linguística experimental

Diana Santos
SINTEF Telecom and Informatics

O que é um «workshop»?

Uma oficina em que todos trabalham?

Uma conferência de pouco vulto?

Um «chá trabalhante»?

Uma loja de ideias?

Um «trabalho de grupo»? «Aula de grupo»...

Utilidade de corpora

- banho de língua
- material de consulta / inspiração / documentação
- confirmador de hipóteses sobre o uso da língua
- descobridor de factos / regularidades
- testador de programas
- desbaratador / complificador de teorias
- testador de teorias sobre a língua
- criador de recursos
- criador de programas

Fundamentos

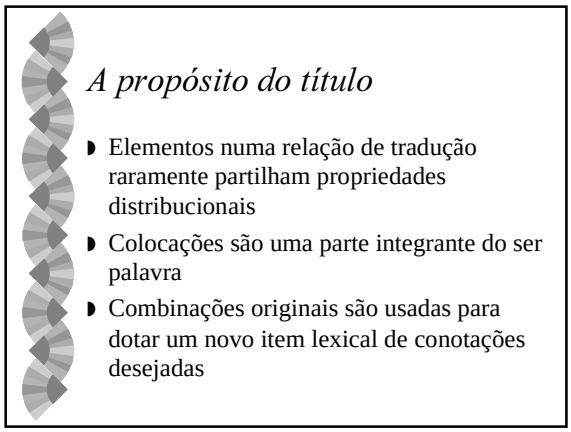
- É preciso **muito** engenho para interrogar um corpus
  - para separar o trigo do joio
  - para pôr as perguntas certas aos interlocutores certos
- Linguística com corpora é uma **conversa** que leva a hipóteses sucessivas, não a uma única interrogação

Fundamentos (cont.)

- um corpus é sempre limitado
  mas aponta no caminho do futuro
- a lei de Zipf é uma constante
  mas os casos raros não são para desprezar
- há que distinguir entre consideração de
  - alguns casos
  - casos acima de um dado limiar de frequência
  - todos os casos

Fundamentos: Avaliação de teorias e de programas

- é complicada
- requer, pelo menos,
  - uma certa unanimidade nos critérios
  - uma descrição rigorosa das experiências e dos resultados
  - recursos partilhados
- dos programas espera-se, além disso, a satisfação de muitos outros requisitos


A propósito do título
Elementos numa relação de tradução raramente partilham propriedades distribucionais
Colocações são uma parte integrante do ser palavra
Combinações originais são usadas para dotar um novo item lexical de conotações desejadas