

Medsensorveiledning, POR2102, våren 2019

Diana Santos

(den kommer først på portugisisk, og så på norsk)

1 Descrição de corpos

Compare os corpos ReLi e CienciaViva

Variantes diferentes, assuntos diferentes e níveis de língua diferentes (resenhas de livros por estudantes brasileiros, publicados na internet; textos de divulgação científica por portugueses, publicados na imprensa local)

Tamanho: 150 mil palavras e 500 mil palavras

2 Expressões regulares

Crie uma expressão regular que aceite números inteiros seguidos de a, b ou c. O zero não é aceitável.

(Exemplos de expressões aceites seriam 2c, 34a, 1003b, etc.)

`[1-9][0-9]*[abc]`

ou

`[1-9][0-9]*[a-c]`

ou

`[1-9][0-9]*(a|b|c)`

3 Expressões regulares no AC/DC

Crie uma expressão regular no AC/DC que obtenha palavras iniciadas por "c" seguidas de um adjetivo terminado por "inho".

`[word="c.*"] [pos="ADJ.*" & word="*.inho"]`

4 Pesquisas

O que é mais verde no corpo Museu da pessoa?

Esta pergunta pode ser interpretada de várias maneiras. Admitindo que é a cor verde como cor e não como indicação de outra coisa,

`[pos="N.*"] [pos="ADJ" & func="N<" & grupo="Verde"]`

Distribuição de lema

dá limão, caminhão e parte

Se for a palavra *verde* (com os seus vários sentidos):

`[pos="N.*"] [pos="ADJ" & func="N<" & lema="verde"]`

Distribuição de lema

dá cinturão

5 Pesquisas

Qual o autor que fala mais de pernas no corpo OBras?

[grupo="Perna"]

Distribuição de autor

É preciso fazer a distribuição relativa

e é CN (Coelho Neto) com .001866

6 Pesquisa semântica

Em que tema do corpo ConDiv se fala mais de doença? E de saúde?

[lema="doença"]

Distribuição por tema

[lema="saúde"]

Distribuição por tema

Como seria de esperar, é no tema da saúde.

7 Pesquisa sobre género

Quem menciona mais o corpo humano no corpo Museu da Pessoa, são homens ou mulheres?

[sema="corpo.*"]

Distribuição de sexo

Fazendo a distribuição relativa, são as mulheres

M: 0.001327945

F: 0.001608915

8 Distribuição

Qual a caracterização mais comum de pessoas no CHAVE? (Lembre-se dos participios passados)

Existe diferença entre as duas variantes?

[lema="pessoa"] (@[pos="ADJ.*"]|@[temcagr="PCP"])

Distribuição dos lemas

físico 1101

jurídico 514

ligar 283

humano 233

Se pusessem outras expressões de procura para pessoas, como por exemplo adicionando homens e mulheres,

[lema="pessoa|homem|mulher"] (@[pos="ADJ.*"]|@[temcagr="PCP"])

Distribuição de lema

físico 1102

jurídico 514

armar 503

público 406
ligar 332

Existe diferença entre as duas variantes? Bastante:

[lema="pessoa|homem|mulher" & variante="BR"] (@[pos="ADJ.*"]|
@[temcagr="PCP"])

aparece "comum" na quinta posição

[lema="pessoa|homem|mulher" & variante="PT"] (@[pos="ADJ.*"]|
@[temcagr="PCP"])

armar 249

forte 221

infectar 203

ligar 191

colectivo 143

9 Pesquisas sintáticas

Obtenha os dois objetos mais comuns do verbo comer no corpo Vercial.

[lema="comer" & pos="V.*"] [pos!="V.*"]* @[func="<ACC"] within s

Distribuição dos lemas

pão 61

carne 35

10 Avaliação de sentidos

Através de uma amostra aleatória de 20 casos no corpo NOBRE, verifique que sentidos consegue obter da

palavra pena como substantivo (que pode significar, em norueguês, medfølelse, fjær, penn, straff, etc, ...)

[lema="pena" & pos="N.*"]

Concordância

Amostra aleatória de 20 linhas.

O aluno deve copiar os 20 casos obtidos e dar uma tradução/classificação a cada...

E eventualmente depois dizer qual a mais frequente e que outros sentidos obteve (fazer um resumo).

11 Emoções

Qual a distribuição do sentimento de vergonha verbal (i.e., como verbo) em termos de pessoa (primeira,

segunda, terceira, do singular e do plural)?

É semelhante ao da globalidade das emoções?

E ao dos verbos em geral?

[sema=".*emo:vergonha.*" & pos="V.*" & pessnum=".*1S.*"] 221 casos

[sema=".*emo:vergonha.*" & pos="V.*" & pessnum=".*2S.*"] 41 casos

[sema=".*emo:vergonha.*" & pos="V.*" & pessnum=".*3S.*"] 1077 casos
[sema=".*emo:vergonha.*" & pos="V.*" & pessnum=".*1P.*"] 19 casos
[sema=".*emo:vergonha.*" & pos="V.*" & pessnum=".*2P.*"] 17 casos
[sema=".*emo:vergonha.*" & pos="V.*" & pessnum=".*3P.*"] 203 casos
Distribuição de pessnum

Fazer o mesmo para a globalidade das emoções
[sema=".*emo:.*" & pos="V.*" & pessnum=".*1P.*"] etc.
9696, 1795, 32495, 740, 322, 5283

E o mesmo para os verbos em geral:
94720, 17419, 546390, 15196, 3292, 104940

O que se pode comentar é que os verbos de emoção são mais frequentes na primeira pessoa do singular e na segunda, e são muito menos frequentes no plural.
Em relação à vergonha comparada com as outras emoções, parece ter mais casos nas terceiras pessoas do que as emoções como um todo.

12 Verbos de dizer
Há mais falas femininas ou masculinas na obra *Dom Casmurro* de Machado de Assis (no Obras)? Sugestão: use o género do sujeito dos verbos de dizer.

[func="SUBJ>.*"] [sema="dizer.*" & obra="Dom_Casmurro"]
Distribuição de genero

Há mais falas masculinas

13 Opinião
Comente o que aprendeu sobre a língua portuguesa nesta cadeira, e o que foi mais surpreendente, o que foi mais gratificante, e o que foi menos interessante. Pode também fazer comentários sobre assuntos que gostaria de ter aprendido ou sugestões de outras questões ou métodos.

Evidentemente que esta pergunta não tem uma resposta padrão. Mas pode ver-se a seriedade e quanto o aluno aprendeu ou quereria ter aprendido.

NORSK:

1 Beskrivelse av korpus
Sammenlikn korpusene ReLi og CienciaViva

Forskjellige varianter, forskjellige emner og ulike språknivå (bokanmeldelser for brasilianske studenter, publisert på internett vs. populariserte tekster om vitenskap for portugisere, publisert i lokalpressen).
Omfang: 150 000 ord og 500 000 ord.

2 Regulære uttrykk

Skriv et regulært uttrykk som kan ta som verdi heltall fulgt av a, b eller c. Kan ikke ta null som verdi. (Eksempler på riktige uttrykk er 2c, 34a, 1003b, osv.)

[1-9][0-9]*[abc]

eller

[1-9][0-9]*[a-c]

eller

[1-9][0-9]*(a|b|c)

3 Regulære uttrykk i AC/DC

Skriv et regulært uttrykk i AC/DC som finner ord som begynner på "e" fulgt av et adjektiv som ender på "inho".

[word="c.*"] [pos="ADJ.*" & word=".*inho"]

4 Søk

Hva er mest grønt (*verde*) i korpuset Museu da pessoa?

Dette spørsmålet kan tolkes på forskjellige måter.

Gitt at det er grønt/verde som farge og ikke i noen annen betydning:

[pos="N.*"] [pos="ADJ" & func="N<" & grupo="Verde"]

Distribuição de lema

gir som resultat *limão, caminhão* og *parte*

Hvis det dreier seg om ordet *verde* (i alle sine betydninger):

[pos="N.*"] [pos="ADJ" & func="N<" & lema="verde"]

Distribuição de lema

gir som resultat *cinturão*

5 Søk

Hvilken forfatter skriver oftest om bein/pernas i korpuset Obras?

[grupo="Perna"]

Distribuição de autor

Det er nødvendig å sette opp en relativ fordeling, og det er CN (Coelho Neto) som vinner med .001866

6 Semantisk søk

I hvilket tema i korpuset ConDiv er det oftest snakk om *doença* (sjukdom)? Og om saúde (helse)?

[lema="doença"]
Distribuição por tema
[lema="saúde"]
Distribuição por tema

Slik en kunne vente, er det i temaet saude.

7 søk etter kjønn
Hvem nevner det menneskelige lege oftest i korpuset Museu de Pessoa, menn eller kvinner?

[sema="corpo.*"]
Distribuição de sexo

Hvis en setter opp en relativ fordeling, er det kvinner.

8 Fordeling
Hva er den vanligste karakteriseringen av personer i CHAVE? (Husk på preteritumpartisippene) Fins det forskjeller mellom de to variantene?

[lema="pessoa"] (@[pos="ADJ.*"]|@[temcagr="PCP"])
Distribuição dos lemas
físico 1101
jurídico 514
ligar 283
humano 233

Hvis en skal bruke andre søkeuttrykk om personer, ved å søke i tillegg på både menn og kvinner,

[lema="pessoa|homem|mulher"] (@[pos="ADJ.*"]|@[temcagr="PCP"])
Distribuição de lema
físico 1102
jurídico 514
armar 503
público 406
ligar 332

Fins det forskjeller mellom de to variantene? Ja:

[lema="pessoa|homem|mulher" & variante="BR"] (@[pos="ADJ.*"]|
@[temcagr="PCP"])
her kommer "comum" i femte posisjon (i stedet for ligar)

[lema="pessoa|homem|mulher" & variante="PT"] (@[pos="ADJ.*"]|
@[temcagr="PCP"])
armar 249
forte 221
infectar 203

ligar 191
colectivo 143

9 Syntaktiske søk

Finn de to vanligste objektene for verbet *comer* (spise) i korpuset Vercial.

[lema="comer" & pos="V.*"] [pos!="V.*"]* @[func="<ACC"] within s

Distribuição dos lemas

pão 61

carne 35

10 Betydninger

Undersøk hvilke betydninger du kan finne av ordet *pena* som substantiv (på norsk 'medfølelse'; 'fjær'; 'penn'; 'straff' osv.) ved hjelp av at tilfeldig utvalg av 20 forekomster i korpuset NOBRE

[lema="pena" & pos="N.*"]

Concordância

Amostra aleatória de 20 linhas.

Studenten skal skrive av de 20 forekomstene, og gi en oversettelse eller klassifisering av hver av dem.

Og eventuelt fortelle hva som er den hyppigste betydningen og hvilke andre som han/hun har funnet (skrive et sammendrag).

11 Emosjoner

Hvordan er verb som uttrykker følelsen *vergonha* (skam) fordelt på person (første person, annen person, tredje person i entall og flertall) i korpuset OBras?

Svarer denne fordelingen til den av alle uttrykk for følelser? Og til den av verb i allmennhet?

[sema=".*emo:vergonha.*" & pos="V.*" & pessnum=".*1S.*"] 221 casos

[sema=".*emo:vergonha.*" & pos="V.*" & pessnum=".*2S.*"] 41 casos

[sema=".*emo:vergonha.*" & pos="V.*" & pessnum=".*3S.*"] 1077 casos

[sema=".*emo:vergonha.*" & pos="V.*" & pessnum=".*1P.*"] 19 casos

[sema=".*emo:vergonha.*" & pos="V.*" & pessnum=".*2P.*"] 17 casos

[sema=".*emo:vergonha.*" & pos="V.*" & pessnum=".*3P.*"] 203 casos

Distribuição de pessnum

Gjør det samme med alle (slags) følelser

[sema=".*emo:.*" & pos="V.*" & pessnum=".*1P.*"] osv.

9696, 1795, 32495, 740, 322, 5283

Og det samme med alle verb:

94720, 17419, 546390, 15196, 3292, 104940

Noe en kan kommentere, er at emosjonsverb forekommer hyppigere i første person entall og i annen person, og at de er meget mindre hyppige i flertall.

Når det gjelder skam sammenliknet med de andre emosjonene, virker det som om de opptrer oftere i tredje person enn (uttrykke for) emosjoner i allmennhet.

12 Ytringsverb

Er det flest kvinner eller menn som snakker i verket *Dom Casmurro* av Machado de Assis (i OBras)?

Hint: ta utgangspunkt i genus til subjektet til ytringsverbene.

```
[func="SUBJ>.*"] [sema="dizer.*" & obra="Dom_Casmurro"]
```

Distribuição de genero

Det er flest menn som snakker.

13 Egne synspunkter

Fortell hva du lærte om portugisisk språk på dette kurset, og si hva som overrasket deg mest, hva som ga deg mest, og hva som var mindre interessant. Du kan også skrive om emner som du gjerne skulle ha lært noe om eller komme med forslag til andre emner eller metoder.

Det er helt klart at dette spørsmålet ikke har noe fast svar, men ut fra svaret kan en se hvor alvorlig studenten har tatt studiene i tillegg til hva studenten har lært eller gjerne skulle ha lært.