

José Mariano Gago – o ministro da língua

Diana Santos

d.s.m.santos@ilos.uio.no



16 de maio de 2018



Uma grande responsabilidade

José Mariano Gago foi, para mim



- um dos meus grandes professores
- um dos meus inspiradores
- um grande amigo e conselheiro
- “o” chefe

O que eu lhe devo não pode, nunca, ser retribuído.



A relação de José Mariano Gago com a língua



Nem todos os presentes saberão, contudo, da importância que José Mariano Gago atribuía à língua e, portanto, ao português como ferramenta:

- Ele viu imediatamente que uma das “armas” que nós tínhamos era a nossa língua, e que uma das formas de combater a subalternização ao inglês era desenvolver o seu estudo e estatuto, na informatização e na sociedade do conhecimento
- Por isso, essa área para ele era estratégica, e foi sempre alvo de conversas e tratados e protocolos com, por exemplo, o Brasil – embora com êxito moderado



A relação de José Mariano Gago com a língua



- Por isso também, tendo as conversas em 1998 com a Microsoft para desenvolver produtos para o português com o apoio do Estado português gorado, não hesitou em lançar uma discussão pública sobre a área, que teve lugar em 1999 com base num documento escrito por mim
- Mesmo (eu) tendo posto quase toda a comunidade académica contra mim (por total falta de qualidades políticas), continuou a investir na área através da criação da Linguatca



A relação de José Mariano Gago com a língua – a nível pessoal



- Era um comunicador excelente
- Tinha línguas afetivas – o português e o francês
- Era muito lido, e apaixonado por literatura (Joyce, por exemplo)
- De acordo com o testemunho de Clara Pinto Correia, aprendeu alemão... para ler Goethe!

A relação de José Mariano Gago com a língua, em duas palavras



- Provavelmente, foi o único ministro da Ciência++ que teve sob o seu pelouro e iniciativa direta, o processamento computacional da língua
- “Pluralismo político”: Usou variados métodos e medidas (a Linguatca é apenas um deles) para que a questão da língua progredisse. Sem me alongar – porque apenas os vi como espectadora
 - o centro de PLN da Microsoft,
 - os acordos CMU-Portugal,
 - área de projetos “Processamento da língua portuguesa” na FCT

- Melhor do que redigir mais uma apresentação, apresento-vos várias formas como foi apresentada ao longo dos tempos, para diversos públicos.
- O nome *Linguateca* apenas surgiu em 2002:
 - antes chamava-se *projeto Processamento Computacional do Português*
 - e depois *Centro de Recursos – distribuído – para o processamento computacional da Língua Portuguesa (CRdLP)*

Atibaia, Brasil, 2000

Three main areas

- Catalogue of NLP/CL of Portuguese
the most visible
- Resource distribution and creation
the heaviest workload
- Evaluation
the most difficult

O que é a Linguateca

- Uma iniciativa para melhorar o processamento do português
 - Disseminação
 - Criação de recursos
 - Avaliação
- Uma organização virtual com quatro pólos
 - Oslo, Braga, Lisboa, Porto, ...
 - Colaboração com outras localizações: Odense, Lisboa, São Carlos, Porto Alegre, ...
- A continuação do projecto *Processamento Computacional do Português*, lançado em 1998, pelo então Ministério da Ciência e da Tecnologia (projecto POSI até 2006)

O que é a Linguateca?

- Um projecto de infraestrutura para o processamento computacional da língua portuguesa
- proposto no seguimento da reflexão em torno da área feita para o Livro Branco (1999)
- é preciso desenvolver recursos para a nossa língua
- é preciso organizar avaliações conjuntas para aumentar a qualidade dos sistemas
- é preciso estabelecer um portal que informa sobre a área

Never heard about *Linguateca*?

- It is a government funded initiative to significantly raise the quality and availability of resources for the **computational processing of Portuguese**
- After an initial plan for discussion by the community (white paper) a network was launched, headed by a small group (Linguateca's Oslo node) at SINTEF ICT (formerly SINTEF Tele og Data)
- This network had as main goal to guarantee that
 - Information was provided and gathered at one place on the Web
 - Resources were made public, maintained, and further developed in connection with the scientific community
 - Evaluation initiatives were launched

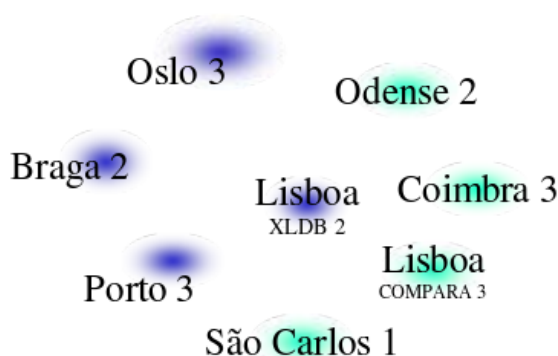
Linguateca, a project for Portuguese

- A distributed resource center for Portuguese language technology

IRE model

- Information
- Resources
- Evaluation

www.linguateca.pt



Context: Linguateca

- **Linguateca** is a network for the computational processing of Portuguese (language) started by Portuguese research ministry after a public discussion in Portugal (1988-1989) about scientific policy.
- It was a distributed initiative with several nodes, the leading one at SINTEF in Oslo, which in addition to resource development had a heavy workload on evaluation venues.
- From 2010 onwards its funding was severely reduced, but the repositories and a subset of the projects has continued to this day, with some new projects happening due to my work at UiO.
- Gramateca, and PANTERA, are just two of them.

A Linguateca vista dos meandros do financiamento

De um ponto de vista político, aquilo em que se tornou a Linguateca teve várias ancoragens e interpretações:

- Uma iniciativa do ministro
- Um projetinho da Agência de Inovação
- Um projeto POSI para o qual a FCCN pediu financiamento
- Um projeto infra-estrutural da FCCN
- Uma iniciativa à qual a FCT/FCCN dá guarida informática



Resultados da procura

pt:166:20:164443:1677:2018
Procura: MI (meset [Bemais"cultural"] [Bemais"cliticaclitica"] *)
Publico de uma concordancia em context: 1
Corpo: Texto de um autor José Mariano Gago v. 3.1
257 ocorrências.
Concordância
Procura: MI (meset [Bemais"cultural"] [Bemais"cliticaclitica"] *)
par=11.19: «Inestimável contribuição para a ciência, tecnologia e a cultura científica em Portugal: Fundação para a Ciência e Tecnologia
par=11.20: A Fundação para a Ciência e Tecnologia (FCT) faz-me esta noite honra, com o privilégio de ser, a norte do antigo ministro
Mariano Gago, dedicado a ser, incondicionalmente, para a ciência, tecnologia e a cultura científica em Portugal.
par=11.26: O Presidente da República, Aníbal Cavaco Silva, escreve, esta noite, sobre uma mensagem de condicionalismo à família do
professor José Mariano Gago, que morreu, afirmando que Portugal perde «uma das personalidades mais marcantes da sua vida científica e
cultural».
par=11.27: «Além de homem de ciência e cultura, José Mariano Gago foi, desde jovem, um cidadão exemplar pelo empenho
demonstrado na defesa intransigente dos valores da liberdade e da democracia. Uma vida...»
par=11.46: Inquirido, sublinhou, entretanto, líder, alguém que teve uma visão para a ciência e a cultura científica do país e, mais do que
isso, a por um público homem, passou, ainda, ao nível de uma «obrigação».
par=11.19: Muitas questões entretanto permanecem. Você, não, uma espécie de para a cultura científica?
par=11.7: José Mariano Gago «colocou no centro da ambição política a sociedade de conhecimento e teve uma preocupação incansável
com a democratização da cultura científica, impulsionando através do lançamento do programa Ciência Viva.
par=11.23: Costa Silva, ainda de «promover, ininterruptamente, para democratização da cultura científica», promovida por Mariano
Gago, considerando o lançamento do programa Ciência Viva.
par=11.7: Além de homem de ciência e cultura, José Mariano Gago foi, desde jovem, um cidadão exemplar pelo empenho
demonstrado na defesa intransigente dos valores da liberdade e da democracia.
par=11.23: Uma e sua morte, Portugal perde uma das personalidades mais marcantes da sua vida científica e cultural, a quem muito
deves, talvez, perguntar de investigação dos seus diversos ramos de saber e do conhecimento.

- Criação de um recurso (em português) sobre Mariano Gago, que também contivesse a sua presença, quer na primeira quer na terceira pessoa, para ser usado pelos intervenientes da área do processamento computacional da língua portuguesa
- Material de avaliação para várias áreas especializadas, algumas delas inovadoras, outras tradicionais. Por exemplo:
 - o estudo dos verbos de relato e sua relação com as emoções
 - a obtenção de redes de relacionamentos a partir de texto
 - o estudo da reputação

Os problemas

É só carregar num botão? Não, todas estas coisas exigem reflexão teórica e anotação humana para decidir o que se pretende e o que se pode obter.

- os verbos *reconhecer*, *esperar*, *apreciar* têm vários contextos em que não são verbos de emoção (ou associados a emoção)
- “relacionamentos”: trabalha com, é colega de, pertenceu ao mesmo governo que, é inimigo ou amigo de, concorda com... ? Em princípio existe um número infinito de relacionamentos possíveis, já para não entrar em conta com mentiras, ou mudanças de estatuto
- reputação, tal como definida correntemente em informática, refere-se a identificadores, sítios da internet, nomes de domínios, ou produtos...

Por exemplo, a RepLab 2013 analisa a reputação de companhias mencionadas no Piador (Twitter)

Na prática, existem muitas outras questões que têm de ser resolvidas (ou pelo menos tratadas) antes...

Por exemplo:

- Classes de emoções
- Identificação de entidades e confirmação da sua semelhança ou diferença
- Como se exprime uma relação em linguagem natural (quais os argumentos, voz ativa e passiva...)

E muitos de nós ficamo-nos por estes problemas...

A questão que quero levantar aqui, é que há um imenso esforço linguístico e documental a fazer no estudo de qualquer língua, e que não se resolve apenas (ou sobretudo) com quantidade. E que as ferramentas de processamento de linguagem natural não são meras ferramentas neutras, mas incluem muitas decisões e escolhas (que deviam ser) fundamentadas.

Emoções no dicionário

inveja [ivéʒɐ]. *s. f.* (Do lat. *invidia*). 1. Sentimento de desgosto e ódio simultâneos provocados pelo bem ou pela felicidade de alguém. *Criouva aquelas intrigas todas por inveja, não podia ver ninguém bem. Sentir, ter +; falar com +.* 2. Sentimento que se traduz pelo desejo de possuir aquilo que outrem possui ou de o igualar ou superar em alguma coisa. = CIÚME, COBIÇA, EMULAÇÃO. *Estar cheio de +; sentir, ter +; meter, causar + a.* 3. Tudo aquilo que provoca um desejo de posse, que é objecto de cobiça. **às invejas**, *loc. adv.*, à pouca. **por inveja**, *loc. adv.*, pelo facto de sentir um misto de desgosto e raiva por aquilo que outra pessoa é ou tem. *Tudo o que ele tem feito ao combado é por inveja, por não suportar o seu sucesso. morder-se de inveja*, ficar enraivecido por alguém ser aquilo que se desejaria ser ou possuir alguma coisa que se gostaria de ter. **morrer de inveja**, o m. que *morder-se de inveja*. **roer-se de inveja**, o m. que *morder-se de inveja*.

invejar [ivéʒár]. *v.* (De *inveja* + suf. *-ar*). 1. Querer para si o que é de outrem; desejar ter o que outra pessoa tem; ter inveja. = COBIÇAR. *Invejava-lhe a inteligência. Todos invejam o seu à-vontade. Invejo esse teu vestido.* 2. Sentir despeito, desgosto porque alguém tem mais, é melhor ou é mais feliz; ter inveja. *Uma amiga que sempre a invejou tentou prejudicá-la.*

gratidão [gratidõw]. *s. f.* (Do lat. tardio *gratitudo*, *-inis*). Reconhecimento efectuado de um benefício recebido; qualidade de quem é grato. ≠ INGRATIDÃO. «*Por isso a civilização deve os mais assinalados serviços ao glorioso infante, que, honrando para sempre a sua memória, honrou igualmente a pátria, que agora lhe presta uma homenagem de gratidão*» (F. DE ALMEIDA, *Inf. de Sagres*, p. 367). «*a clientela que, a princípio, o recebera sem agrado e quase com desrespeito, tributava-lhe agora, na generalidade, enorme gratidão.*» (M. G. AZAMBUJA, *Barbara*, p. 20).

Dicionário da Academia

A necessidade da documentação: pelo punho de JMG

Em 2014-05-20 15:27, José Mariano Gago escreveu:



- > Dianinha
- > Tens não só o Direito, mas (se me
- > permites) o dever de o fazer quanto
- > antes.
- > Podes citar-me como entenderes,
- > evidentemente.
- > Acho que deves citar os episódios que
- > te pareçam relevantes com o detalhe
- > necessário e sem reservas.
- > Um texto desta natureza é fundamental
- > como referência cultural e como recurso
- > histórico.



Obrigada!

