

Pareceres sobre *Linguateca*: Relatório relativo ao período 2000-2003

Foi pedido, por correio electrónico, a vários especialistas na área ou áreas afins, *que não tivessem colaboração institucional com a Linguateca*, que produzissem um parecer em dois parágrafos sobre a actividade do nosso projecto.

Agradecemos a resposta da quase totalidade das pessoas a quem nos dirigimos, e publicamos os pareceres pela ordem por que nos chegaram.

Eugénio Oliveira, FEUP

O Relatório sobre a Linguateca, relativo ao período 2000-2003 apresenta de uma forma clara, estruturada e sumariada, as actividades e os resultados do trabalho realizado pela equipa de investigação do Projecto "Centro de Recursos Distribuído para o Processamento Computacional da Língua Portuguesa".

O Relatório permite uma rápida familiarização com os temas, objectivos e produtos, relativos ao trabalho efectuado, o qual reputo de importância para o público alvo referido no Relatório.

Maria Francisca Xavier, FLUNL

A *Linguateca* tem efectivamente funcionado desde a sua criação como um **centro de recursos para a língua portuguesa**, que se tem revelado desde o início uma importante força congregadora dos investigadores interessados no processamento computacional da língua portuguesa, nomeadamente quanto a troca de experiências, a divulgação de informação actualizada, a disponibilização de recursos e de prestação de serviços, a preparação de encontros e workshops, actuando de uma forma totalmente aberta que transmite entusiasmo e confiança tanto aos seus parceiros e formandos como aos próprios clientes. Como uma prova do grande dinamismo e vitalidade da Linguateca será suficiente percorrer a extensa lista de "produção escrita", que teve início apenas em 1999, e que se encontra disponível na página 'online'.

A Linguateca é um centro indispensável para continuar a contribuir para o debate científico das diversas vertentes do seu domínio e para a rentabilização dos recursos já existentes ou em desenvolvimento numa área que se tornou vital para a sociedade actual.

Jorge Baptista, Universidade do Algarve

A *LINGUATECA*, liderada pela Doutora Diana Santos, tem desempenhado desde a sua criação um importantíssimo papel no desenvolvimento do Processamento Computacional do Português, quer na sua actividade de centro de criação e difusão de recursos, quer na promoção de diversas iniciativas, entre as quais avulta a organização de um processo de avaliação conjunta de sistemas e de recursos de processamento de linguagem natural, especificamente desenvolvidos para a nossa Língua. Neste sentido, a *LINGUATECA* conseguiu mobilizar e contar com a participação activa de toda a comunidade científica do domínio (dos dois lados do Atlântico), o que é um feito praticamente ímpar no panorama científico internacional.

A qualidade e as dimensões dos recursos construídos, bem como os inestimáveis serviços de apoio à comunidade científica, a seriedade e o dinamismo que revela na consecução dos seus objectivos e as perspectivas de projectos futuros, de que dá conta o relatório relativo ao período de 2000-2003, tornam a *LINGUATECA*, indiscutivelmente, um importante factor de desenvolvimento global do Processamento Computacional do Português, área científica cujo valor estratégico, apesar já amplamente reconhecido, nunca será de menos realçar.

Isabel Trancoso, Luis Oliveira, Nuno Mamede, IST

Consideramos bastante positivo o trabalho desenvolvido pela Linguateca nos seus três eixos de actividade ao longo destes últimos anos. Pensamos que forneceu uma contribuição muito importante para o processamento da língua portuguesa e merece sem qualquer dúvida ser continuado.

Gostaríamos também de destacar o envolvimento da Dr. Diana Santos nos trabalhos que levaram ao estabelecimento em Fevereiro de 2002 do protocolo que previa a cooperação luso-brasileira em matéria de redes electrónicas, processamento da língua portuguesa e promoção de conteúdos digitais. Infelizmente, o protocolo, apesar de assinado por dois ministros da Ciência e Tecnologia, não avançou. Na nossa opinião, é deveras de lamentar, pois seria uma excelente forma de formalizar a expansão de pólos da Linguateca no Brasil.

A principal crítica que o relatório nos sugere é a da inexistência de um conselho consultivo que se pronuncie sobre a criação e a localização de novos pólos, sobre a aquisição de novos recursos e sobre novas actividades a desenvolver. Consideramos que no futuro se deveria considerar a constituição deste órgão de carácter consultivo.

Violeta Quental, PUC-Rio de Janeiro

Tive a grata oportunidade de ler o **Relatório relativo ao período 2000-2003**, de Diana Santos, sobre as actividades da *Linguateca* e gostaria de testemunhar a importância do trabalho que sua equipe vem desenvolvendo para a comunidade de pesquisa em Linguística Computacional de língua portuguesa e para o público em geral. Venho acompanhando o projeto de criação e manutenção de um repositório de recursos e ferramentas, bibliografia, fórum de discussão e avaliação da área de Processamento Computacional do Português já há muitos anos, e, junto com meus colegas e alunos, tenho utilizado freqüentemente a *Linguateca* para pesquisa e actividades didáticas. A preocupação do projeto com a disseminação de conhecimento e com a disponibilização de recursos e ferramentas, bem como com a avaliação desses mesmos materiais foi canalizada para esse *site* com competência extrema.

Como demonstra o relatório, a actividade da equipe está-se traduzindo em cada dia mais e melhores ferramentas disponibilizadas e em um engajamento crescente da comunidade de pesquisa na troca de conhecimento e na avaliação da pesquisa na área. Esse movimento de transformação de uma série de pessoas e grupos isolados trabalhando em processamento computacional do português em uma comunidade de pesquisa e colaboração deve-se, essencialmente, à equipe da *Linguateca* e, apenas por esse motivo, já seria seu trabalho plenamente justificado. Acresce-se a esse fator agregador a qualidade acadêmica, a seriedade e cuidado técnico com que têm sido tratados os

diversos serviços prestados pela *Linguateca*. A comunidade científica da área só tem a agradecer à Diana Santos e aos demais membros da equipe da *Linguateca* por seu trabalho e a apoiar sua continuidade e crescimento.

Graça Nunes, Universidade de São Carlos

O serviço prestado pela Linguateca à comunidade científica de Processamento Computacional do Português tem alterado para melhor a forma de se fazer pesquisa de qualidade nessa área. Os pesquisadores podem contar com recursos cuidadosa e especialmente preparados segundo procedimentos estabelecidos na literatura internacional. Uma maior integração e cooperação entre equipes do Brasil e Portugal e a possibilidade de encurtar a distância dos trabalhos envolvendo o Português daqueles que envolvem o Inglês constituem dois dos muitos efeitos positivos dessa iniciativa. Como resultado, a impressão que se tem é que toda a comunidade já se sente parte da equipe da Linguateca.

Graciete Teixeira, Porto Editora

O relatório relativo ao período de 2000-2003 da Linguateca apresenta as actividades da mesma, sendo manifesto o cumprimento dos seus objectivos.

A disponibilização de recursos na rede revela-se da maior utilidade para todos, principalmente para os investigadores empresariais, dado que os meios disponíveis nas Universidades são acessíveis quase só à comunidade científica. Por experiência própria, a Porto Editora conhece a dificuldade em obter recursos devido à falta de modelo comercial por parte das Universidades. No entanto, os nossos produtos apenas utilizariam alguns dos resultados obtidos pela nossa investigação a partir desses recursos, que, por outro lado, são disponibilizados gratuitamente para investigação puramente científica. Deste modo, a actividade da Linguateca veio suprir uma lacuna importante no processamento computacional do Português e o trabalho desenvolvido destaca-se pelo interesse que representa para o estudo das estruturas internas da nossa língua e para a própria consolidação da Língua Portuguesa.

Lucia Rino, UFSC

O relatório sobre a Linguateca, relativo ao período 2000-2003, está muito claro e bem apresentado, demonstrando o imenso avanço e a importante contribuição para as comunidades interessadas no processamento da língua portuguesa.

Vale ressaltar que, embora o foco da Linguateca seja o processamento computacional do português, os recursos ora agregados para a disseminação, construção e disponibilização de recursos dessa língua têm uma abrangência muito maior, por permitirem estudos dos mais variados. Como exemplo, podemos citar aqueles estritamente lingüísticos, contrapondo-se aqui ao foco computacional da Linguateca. Desse modo, são várias as comunidades que podem se beneficiar desta, envolvendo não só a diretamente relacionada ao processamento automático da língua portuguesa, como também as comunidades de lingüística, comunicação e filosofia, dentre outras.

Renata Vieira, Unisinos

O projeto Linguateca apresenta como resultados a organização e disponibilização de uma série de recursos para o processamento da língua Portuguesa. O resultado desse trabalho é muito útil tanto para o ensino de Lingüística Computacional como para o ensino de Lingüística e da Língua Portuguesa. Pesquisas realizadas nessas áreas podem também se beneficiar desses recursos. Esse trabalho de divulgação e organização dos recursos em único local de acesso é de grande auxílio àqueles que desenvolvem ou querem iniciar estudos nessas áreas. É interessante constatar que esses recursos tem chamado a atenção não só de brasileiros e portugueses, mas como de interessados de vários outros países.

Paralelamente ao trabalho prático de organização de recursos há um trabalho forte de divulgação e de pesquisa e desenvolvimento de recursos, como pode ser observado pelas publicações e apresentações realizadas no período. Além disso o projeto tem estimulado a discussão na comunidade das questões relativas à avaliação de trabalhos ligados ao PLN .

Tony Berber Sardinha, PUC/SP

No meu entender, a Linguateca é um projeto de grande impacto, que tem tido um papel de grande destaque junto à comunidade de usuários de corpora e de recursos para análise informatizada da língua portuguesa. Ainda segundo minha avaliação pessoal, o maior destaque do portal refere-se ao acesso a corpora eletrônicos por meio de concordanceadores *online*. O sistema de busca por palavras e demais expressões é bastante simples e acessível até mesmo a usuários com pouco conhecimento de informática. Um ponto que talvez possa ser enfatizado no futuro seja o de proporcionar ferramentas com o mesmo nível de simplificação para outras tarefas, tais como etiquetagem morfosintática e alinhamento de textos, para permitir um alcance a um número maior de usuários, entre eles alunos de graduação de Letras, Lingüística e Lingüística Aplicada (vide, por exemplo, a seção de ferramentas em <http://lael.pucsp.br/corpora>).

Isabel Hub Faria, UL

Estive a ler o seu relatório que acho muito informativo e que mostra bem o desenvolvimento que você e os seus colegas têm conseguido imprimir aos vossos projectos. Para alguém que não esteja completamente familiarizada com os conteúdos (que é o meu caso), o relatório é pouco explícito. A introdução mais pormenorizada das potencialidades dos diversos materiais, do ponto de vista do potencial utilizador, poderia constituir uma vantagem. Os objectivos globais do vosso projecto parecem-me dispersos pelos projectos constituintes, mas valia a pena caracterizá-los inequivocamente, podendo constituir um incentivo à sua utilização e ao próprio alargamento da Linguateca. Penso que o esforço que tem posto no seu trabalho tem sido notável ao longo dos anos. Parabéns!

Oswaldo Oliveira Jr, NILC

O relatório de atividades do Projeto Linguateca, para o período de 2000 a 2003, traz as principais diretrizes que norteiam o projeto assim como os resultados significativos obtidos. Da leitura do relatório e da minha experiência acompanhando usuários das ferramentas e recursos da Linguateca, resalto a importância do Projeto em dois

contextos principais: 1) estratégico para a inserção da língua portuguesa no mundo digital; 2) da consecução dos objetivos mais específicos do projeto.

Quanto ao valor estratégico do Projeto Linguateca, enfatizo a consolidação da concepção de que a manutenção de um centro de recursos lingüísticos e lingüístico-computacionais é a forma mais eficaz de atingir a comunidade de uma determinada língua, principalmente com vistas a permitir um volume crescente de conteúdos digitais na rede WWW (World Wide Web) daquela língua. No que concerne à consecução dos objetivos, quais sejam os de disponibilizar e manter recursos lingüístico-computacionais para o processamento automático e estudos lingüísticos do português, o relatório mostra claramente a evolução dos recursos e ferramentas no período de 2000 a 2003. O Projeto Linguateca vem permitindo uma integração sem precedentes da comunidade de processamento automático, e fazendo com que a língua portuguesa seja privilegiada em comparação com outras línguas com grande número de falantes no que diz respeito à infra-estrutura lingüístico-computacional. São dados merecedores de atenção o aumento vertiginoso no número de consultas ao projeto AC/DC, de um valor incipiente em 2000 para aproximadamente 4000 consultas mensais em 2003, assim como o número de consultas ao Portal Linguateca.

Chama também a atenção a participação da comunidade brasileira, tanto no que concerne ao fornecimento de recursos lingüístico-computacionais para a variante do português do Brasil quanto aos acessos e utilização dos recursos em geral. Isto vem sendo feito sem o envolvimento oficial de órgãos de fomento brasileiros, mas serviu de subsídio para um tratado de cooperação entre os governos brasileiro e português, assinado em Fevereiro de 2002, para o avanço do processamento automático da língua portuguesa.

Concluindo, espero que a iniciativa da Linguateca seja continuada e estendida, quiçá com a participação de outros países lusófonos.