

Escola de Linguística de Outono 2022

Rolezinho Linguístico

Olá! Bem-vindo à segunda atividade olímpica da décima primeira Escola de Linguística de Outono: o Rolezinho Linguístico!

Esta atividade é em grupo e vale 90 dos 360 pontos possíveis durante a ELO. Para esta atividade, você deverá desenvolver uma investigação empírica a partir de uma pergunta de pesquisa da Linguística.

Cada grupo tem 4 participantes, sorteados entre os participantes da ELO que confirmaram participação nesta atividade. Cada grupo trabalhará em um dos seis campos de pesquisa possíveis para este rolezinho: **Olfa Youssef, Nadim, Fihri, Fakhr-un-Nisa, Jazari, Baqir**. Cada campo tem um orientador, dentre professores e pesquisadores de instituições federais, e será trabalhado por dois grupos, a serem escolhidos pelos times.

Para executá-la, você tem duas semanas: de 09 a 21 de maio de 2022. Para isso, você poderá contar com a ajuda e feedback de seu(s) orientador(es), bem como de toda bibliografia confiável que você tiver disponível. O calendário da atividade é o seguinte:

- *Qui, 05 mai:* **Sorteio dos times**, anunciado no grupo dos participantes.
- *Sab, 07 mai:* **Escolha dos campos** de pesquisa por cada time, [nesta tabela](#).
- *Seg, 09 mai (19:15):* **Reunião inicial**: Apresentação do cronograma e conversa inicial com os orientadores de cada campo.
- *Ao longo das duas semanas:* Preparação dos experimentos, coleta dos dados, análise dos dados, preparação da apresentação.
- *Ao final da segunda semana:* **Encontro com os orientadores** para análise de dados e treinamento da apresentação.
- *Sex, 20 mai:* **Gravação da apresentação**: preparação e envio da gravação da apresentação de cada time.
- *Sab, 21 mai (14:00):* **Apresentação dos trabalhos**: Pitch inicial e sessão de pôsteres virtuais + sessão de perguntas e comentários.

No encontro inicial, os orientadores ajudarão a avaliar, orientar e esclarecer pontos obscuros, bem como a delimitar o objetivo específico e os métodos de coleta e análise. Ao longo das semanas, eles ficarão disponíveis, através de um grupo de mensagens específico, para ajudar na coleta e análise, bem como assistir à apresentação previamente e ajudar a melhorá-la. Os encontros intermediários entre grupos e orientadores serão marcados em cada caso. É permitido o uso de qualquer material teórico, ficando a cargo dos estudantes providenciarem-no, caso julguem necessário.

O trabalho final será apresentado em três passos:

- Um **pitch oral**, em que cada time apresentará para os demais, em 90 segundos, de forma criativa, os pontos essenciais de sua pesquisa.
- Uma **sessão virtual de apresentações**, em que cada time poderá preparar uma apresentação mais completa do seu trabalho, no formato de uma apresentação em vídeo de até 7 minutos, com as seguintes restrições:
 - O apoio visual pode se dar na forma de cenas de animação ou de uma apresentação de slides.
 - No caso de slides, pode-se utilizar no máximo 5 páginas de slide (+ slide de capa + slide com bibliografia essencial);
 - Cada slide ou cena pode ter no máximo 15 palavras de texto escrito, descontando-se preposições e conjunções. Os dados coletados e os exemplos do experimento, quando escritos, não contam para este limite. Números e símbolos matemáticos também não contam.
 - De uma forma geral, recomendamos dar preferência a esquemas visuais, imagens e gráficos, na apresentação.
- Uma **sessão de perguntas e respostas**, em que os grupos ficarão disponíveis em diferentes canais de voz e vídeo, tirando dúvidas dos jurados, dos demais times e de outros possíveis visitantes. Para esta seção, é recomendado ter em mãos slides adicionais que expliquem o detalhamento do seu trabalho.

Cada trabalho receberá uma pontuação, atribuída da seguinte forma:

- **60 pontos** de avaliação por um júri especialista, entre pesquisadores e alunos de pós-graduação. Cada um deles dará uma nota de 0 a 10, seguindo a grade de Critérios de Avaliação do Rolezinho Linguístico.
- **20 pontos** de avaliação pelos outros grupos. Cada grupo deve avaliar todos os outros trabalhos sendo apresentados no dia, exceto o seu, seguindo os mesmos critérios do júri especialista.
- **10 pontos** de auto avaliação. Cada grupo deve, ao final do processo, reunir-se e decidir conjuntamente uma nota de 0 a 10 para cada um de seus membros.

Os demais detalhes serão conversados com os respectivos orientadores.

Boas investigações!

Processamento de palavras em português brasileiro

O que acontece em nossa mente quando lemos ou ouvimos uma palavra? Será que para entendermos uma palavra como *roseira*, primeiro precisamos mentalmente acessar o significado de *rosa* para só então entendermos que *ros-eira* é a planta que gera rosas? Ou será que todas as palavras estão organizadas em nossa mente separadamente? Sabemos que o processamento de palavras é muito rápido e intuitivo. Entretanto, entender quais os fatores e como esses fatores influenciam nesse processamento tem sido objeto de estudo da Psicolinguística e da Neurociência da Linguagem.

Por exemplo, será que o reconhecimento da palavra *madeira* se dá da mesma maneira que o da palavra *roseira*? Se olharmos esses dois exemplos mais de perto, verificamos que apesar da terminação em *-eira* estar presente nas duas palavras, há características específicas e distintas entre esses elementos – em *roseira* facilmente identificamos o *-eira* como um sufixo que forma uma palavra derivada *ros(a)eira*. No caso de *madeira*, não há uma derivação, *-eira* é a terminação da palavra monomorfêmica e indivisível *madeira*.

Uma das subáreas da Psicolinguística Experimental busca justamente entender e explicar esse processamento com bases empíricas (experimentos linguísticos). Um dos temas de interesse é investigar se ao acessarmos e processarmos uma palavra levamos em conta as características morfológicas, como em *roseira*, ou, se a informação relativa à raiz e aos morfemas categorizadores não influencia o tempo e a dificuldade de processamento e, então, processaríamos a palavra de maneira inteira, como em *madeira*.

Neste projeto, vocês devem focar em um fenômeno morfológico e/ou semântico e investigar: **como os falantes de língua portuguesa acessam e processam palavras?** Para isso, vocês deverão escolher um fenômeno linguístico que possa suscitar questões interessantes e construir e executar um **experimento linguístico** para testar a hipótese aventada.

Vocês vão utilizar a técnica experimental de PRIMING (França et al, 2015), em que duas palavras são apresentadas em sequência (por exemplo, primeiro, *rosa* e depois, *roseira*) e se investiga o efeito de facilitação ou não que a primeira palavra (*prime*) tem sobre a outra (*target*). O resultado (a facilitação ou não) é entendido como reflexo do acesso lexical. O grupo poderá escolher uma estrutura, como, por exemplo, as sugeridas no quadro abaixo:

Relação	Prime	Target
Morfológica	LEITURA	leitor
Semântica	LEITURA	livro
Fonológica	LEITURA	leite
Composicional	ARCO	arco-íris
Derivação	GLOBAL	globalização
Não há	LEITOR	carro

Vocês devem levar em consideração os achados das teorias de **processamento morfológico e acesso lexical** para construir e executar o experimento, seguindo os seguintes passos:

- Definir um *design* experimental para testar o acesso e/ou o processamento de palavras;
- O *design* experimental deve conter, além do fenômeno linguístico escolhido dentro da área de Processamento de Palavras, a(s) variável(is) independente(s) que será(ão) manipulada(s), a técnica experimental que será utilizada, a(s) variável(is) dependente que será(ão) aferida(s) e as condições experimentais;
- Elencar os objetivos e as hipóteses com base na literatura existente;
- Descrever os resultados, analisá-los e interpretá-los.

Espera aí, como foi mesmo que você disse?

Temos percebido a necessidade de tornar acessível para um público não especializado aquilo que é desenvolvido, criado e pensado a partir de métodos científicos. A importância desse movimento que cientistas, pesquisadores e jornalistas têm tomado cresce principalmente por conta da proliferação de discursos anti e pseudo-científicos, negacionistas ou baseados em FakeNews, os quais resultam na manipulação da opinião pública, fomentando a ausência do pensamento crítico frente às informações.

Esse movimento é conhecido como **Divulgação Científica (DC)**. As pessoas envolvidas em ações de DC têm o objetivo de popularizar e acessibilizar a informação a partir de evidências comprovadas cientificamente. Entretanto, você deve saber que, muitas vezes, os suportes/veículos, objetivos, procedimentos, resultados e conclusões que são transmitidos por pesquisadores e cientistas podem não ser inteligíveis para um público não especializado.

Na maioria das vezes, os agentes científicos apagam os sujeitos para criarem um sentido de objetividade, distanciam-se dos leitores e de seus objetos de pesquisa, impessoalizam *o que é dito* por meio do *como é dito*. Em contrapartida, os divulgadores científicos preocupam-se em trazer o público para dentro de seus textos, deixando evidências de que o *outro* faz parte de seu projeto de texto. Para tanto, fazem uso de certos recursos linguísticos, dos quais destacamos três:

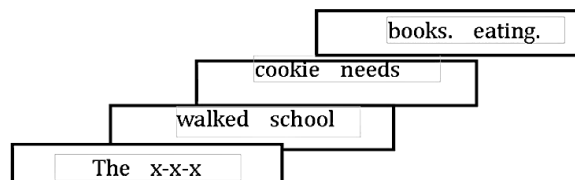
- *Expansão*: Utilização de expansão ou inclusão de informações, com o objetivo de proporcionar os significados conceituais necessários para lograr a efetiva participação cognitiva e comunicativa do leitor;
- *Redução*: Através de certas estratégias linguísticas, como a condensação ou mesmo a eliminação completa, suprime-se dada informação científica, mantendo conceitos de relevância cognitiva e comunicativa;
- *Variação*: Procedimento caracterizado a partir de certas estratégias discursivas de ordem lexical, semântica ou mesmo de registro (entre termos e conceitos especializados e vocabulário corrente) ocorridas durante o processo de reformulação do texto científico para o texto de divulgação.

Essas são **estratégias de interlocução** comumente utilizadas em textos escritos, e pouco se sabe sobre aquelas que são utilizadas quando o suporte informacional é do registro oral. Nesse sentido, a pergunta de nossa pesquisa é: **Quais estratégias de interlocução são utilizadas oralmente por especialistas no contexto de Divulgação Científica?** Para responder essa pergunta, siga os passos:

- a. Escolher uma das cinco temporadas do *Corpus oral de Divulgação Científica*;
- b. Verificar se há estratégias de Expansão, Redução e/ou Variação nas respostas dos especialistas;
- c. Investigar se há outras estratégias de interlocução que podem ser caracterizadas a partir do registro oral;
- d. Verificar se as informações de perfil dos respondentes (idade, sexo, formação e área de atuação) possuem alguma relação com as estratégias de interlocução levantadas nos itens “a” e “b”.

Treinamento linguístico com a *Maze Task* para a aprendizagem de clíticos em próclises

Você já ouviu falar na *Maze Task* ou Tarefa Labirinto? É um **método experimental** usado na Psicolinguística, que tem por objetivo medir quanto tempo as pessoas levam para ler e processar diferentes elementos linguísticos inseridos em uma sentença. Veja como a tarefa funciona: o participante lê frases em um computador, palavra por palavra ou parte a parte, no seu próprio ritmo de leitura. No entanto, após a apresentação da primeira palavra, a tarefa labirinto oferece duas opções de palavras, das quais apenas uma pode completar corretamente a sequência sendo formada. Os participantes são instruídos a escolher rapidamente a opção correta em cada um dos pares apresentados, utilizando, para isso, as teclas especificadas no teclado do computador. Por exemplo, a frase “The school needs books” pode ser apresentada como na ilustração acima (ENKIN; FORSTER, 2014), onde cada quadro representa uma tela que seria apresentada aos participantes.



Recentemente, alguns pesquisadores vêm explorando com sucesso a possibilidade de usar a tarefa labirinto não somente para testagem, mas como uma ferramenta pedagógica na **aprendizagem de estruturas linguísticas** (ENKIN; FORSTER, 2014; OLIVEIRA et al., 2020). De modo geral, os resultados sugerem que os participantes expostos a estruturas específicas, em sessões práticas com a tarefa labirinto, apresentaram melhoras tanto em tarefas que envolvem conhecimento implícito, mensurado em tempos de reação, quanto em tarefas que envolvem conhecimento explícito, mensurado em julgamentos de aceitabilidade. O potencial de aprendizagem está no fato de a tarefa demandar um processamento incremental e permitir a oferta de um feedback imediato. Funciona assim: a instrução formal cria uma memória explícita sobre os aspectos gramaticais e a tarefa labirinto pode oferecer a prática extensiva necessária para converter tal conhecimento em memória implícita.

A partir dessas considerações, para a realização deste estudo, sugerimos um experimento para treinamento linguístico da aprendizagem de clíticos em próclises com a *Maze Task*, com o intuito de verificar: **essa ferramenta é eficaz para o treino e consequente aprendizagem implícita dessas estruturas gramaticais?** Em relação à estrutura-alvo, pronomes clíticos são pronomes pessoais de uma só sílaba (como “o”, “me”, “lo”, “lhe”, “se”) que não têm acentuação própria e que, por isso, dependem do acento da palavra que está imediatamente antes ou depois, geralmente um verbo. Para coleta e análise dos dados, sugerimos que sigam os seguintes passos:

- Recrutar um grupo de participantes para o grupo experimental e realizar um experimento do paradigma Pré-teste + Treino + Pós-teste, em que 1) os testes sejam julgamento de aceitabilidade OU self-paced reading OU cloze e 2) o treino consista em uma sessão com a *Maze Task* focando na aprendizagem implícita de clíticos em próclises;
- Recrutar um grupo de participantes para o grupo controle a realizar o Pré-teste + Treino + Pós-teste, em que os testes sejam os mesmos utilizados pelo grupo experimental e não haja treino;
- Analisar os dados, pensando: existem diferenças estatisticamente significativas considerando-se as médias entre o pré-teste e o pós-teste dos grupos experimental (treino com *Maze Task*) e controle, evidenciando a aprendizagem da estrutura alvo.

“Enfim, a hipocrisia”: quando a teoria, na prática, é outra teoria.

Na época das eleições, muitos candidatos se apresentam para o pleito como os “*outsiders*” da política, mesmo sendo agentes que fazem política de longa data. Entretanto, este não é um fenômeno isolado das campanhas eleitorais, pois pode ser observado em muitos outros lugares, como, por exemplo, quando o investidor (talvez seu primo) “rico” tem o apartamento leiloado por dívidas; ou quando a *black friday* vira *black fraude* com todos os produtos pela metade do dobro; ou quando um religioso defende regras de uma vida correta e íntegra, mas é flagrado em condutas antiéticas, imorais e impróprias; ou ainda quando uma autoridade diz defender a democracia, mas sempre a afronta em suas atitudes. É curioso que, mesmo com esses comportamentos díspares, parte considerável da sociedade ainda acolhe a pessoa hipócrita, justificando suas ações contraditórias com desculpas esfarrapadas, como meros “mal entendidos”.

Isso acontece porque cada falante constrói **uma imagem de si na sua fala pública, um Ethos**, uma representação pessoal conhecida desde a clássica Retórica de *Aristóteles*, mas que tem ganhado uma nova roupagem nos estudos discursivos contemporâneos de *Dominique Maingueneau* e outros. Na língua portuguesa falada no Brasil, o descompasso hipócrita entre o que se diz e o que se faz cristalizou-se em expressões e ditados populares como, por exemplo, “casa de ferreiro, espeto de pau”; “pimenta nos olhos dos outros é refresco”; “Por fora bela viola, por dentro pão bolorento”; ou o famoso “promessa é dívida, hein?!” para tentar garantir que o ato de fala possa ser realizado. Como diria *François La Rochefoucauld*: “A hipocrisia é uma homenagem que o vício presta à virtude”.

Do ponto de vista científico, a categoria do *ethos* mobiliza uma rede conceitual que ultrapassa o falar de si mesmo (autobiografia ou autoavaliação, por exemplo) e também não está limitada às dicotomias “sinceridade/dissimulação” ou “verdade/mentira”, mas atravessa várias áreas de estudo, como a enunciação e a argumentação, por exemplo, visando mostrar as **estratégias pelas quais um locutor atua** na afetividade do seu interlocutor para construir uma imagem psicológica/sociológica sobre si próprio.

Diante disso, nesta proposta vocês devem investigar: **como se caracteriza o ethos construído por discursos socialmente entendidos como “hipócritas”?** Para isso, vocês precisam selecionar falas públicas de uma pessoa ou entidade que construa um ethos diferente do que se observa na realidade. O corpus de pesquisa deve ser delimitado de textos orais ou escritos amplamente veiculados, especialmente em questionamentos de sua imagem pública (texto-resposta / react). Os passos sugeridos são os seguintes:

- a. Escolher um dos seguintes domínios discursivos: saúde e bem-estar (por exemplo, blogueiro *fitness*), marketing (p. ex., propaganda de *pick up*), ou financeiro (p. ex., *coaches* sem experiência);
- b. Coletar um *corpus* de pesquisa público (não resultado de interações em contexto privado), para a análise do Ethos construído por um locutor relevante para o domínio discursivo escolhido;
- c. Estabelecer os objetivos e as hipóteses com base na literatura existente;
- d. Descrever marcas linguísticas (pronomes, expressões, tipologias textuais, atos de fala, entre outras), para evidenciar uma linha de interpretação sobre as estratégias discursivas que constroem o Ethos analisado, apresentando respostas para a pergunta de pesquisa.

“Azul/morto/X de fome”? O processo de intensificação por meios não tradicionais

“Nossa, ele me deixou vermelha de vergonha!”, “Eu tô morta de preguiça!”, “minhas pernas estão ficando roxas de frio”, “morto de fome”. O que esses falantes estão expressando ao usar essas sentenças? É possível que a pele sofra alteração corporal ocasionada por uma sensação/sentimento? Por que usar “morto(a) de”? Outras palavras podem ser utilizadas, em vez de “morto(a)” ou de uma cor, em sentenças como essas? Procurando responder a essas perguntas, pesquisas a respeito dessas expressões (cor/“morto” + de) identificaram que elas (também) são utilizadas para intensificar!

A intensificação é definida por Vieira e Machado Vieira (2008) como um “processo cognitivo avaliativo do mundo, [...] empregado para indicar que a dimensão ou a intensidade de dado elemento ultrapassa os limites do que se concebe como relativamente normal/neutro a ele”. Para entendermos melhor, vejamos alguns outros exemplos:

1. “Fiquei tristinha na hora do almoço, e no fim da tarde recebi um feedback que me deixou **vermelha de alegria**. Têm sido bons tempos”
2. “sai da academia debaixo de chuva **morto de fome** fui direto no bobs comer um sanduiche nada saudavel cheio de sodio com o MEU DINHEIRO porque estava morrendo de vontade”.

Em sentenças similares a (1) e (2), cores ou o adjetivo “morto” – acompanhados pela preposição “de”, localizada imediatamente posposta a eles –, por exemplo, são utilizados para intensificar o que é apresentado após tal preposição: nesses casos, “alegria” e “fome”, respectivamente. Esses usos parecem ser fruto de extensão semântica, de modo que se convencionaram como expressões intensificadoras depois de terem suas características descritivas neutralizadas, isto é, nesses contextos, abandonaram a função descritiva e assumiram a intensificadora.

A proposta de pesquisa é analisar sentenças similares às apresentadas aqui, além de outras expressões intensificadoras atípicas em que uma unidade lexical seja acompanhada pela preposição “de” e, juntas, intensifiquem o que vier depois. Mais precisamente, a questão é: **Qual é a configuração sintática e semântica de algumas dessas expressões intensificadoras?** Para alcançar a resposta, encorajamos os participantes a executar os seguintes passos:

- a. Escolher uma expressão intensificadora atípica estruturada da seguinte forma: X de (intensificador) Y (intensificado);
- b. Delimitar o viés da pesquisa (sincrônico ou diacrônico) e coletar dados (em bancos de textos *on-line*, como o [Corpus do Português](#), ou em um *corpus* próprio, criado a partir da coleta de dados em redes sociais, por exemplo);
- c. Analisar os dados pensando: Quais pessoas do discurso (1a, 2a ou 3a do singular/plural) usam essa expressão? Esses sujeitos são animados ou inanimados? Quais verbos aparecem com mais frequência nessas sentenças? Os itens intensificados são positivos ou negativos? Qual(is) é(são) a(s) temática(s) das sentenças em que essa expressão intensificadora é usada?;
- d. Apresentar respostas para a pergunta de pesquisa, a partir das possíveis regularidades observadas.

Campo Baqir

(Taha Baqir)

Orientadores:

Renato Costa (Turing USP - SP)

Daniel Frozi (Turing USP - SP)

João Pedro Gomes (Turing USP - SP)

Rian Fernandes (Turing USP - SP)

Remoção de stopwords e lemma(tização) para textos jornalísticos

Um dos grandes questionamentos e focos de NLP (Processamento de Linguagem Natural) é como fazer com que a linguagem humana – diversificada em várias expressões, gírias e entonações – seja interpretada e processada por máquinas e computadores que só interpretam números. De maneira geral, contudo, pode-se apontar que a amenização dessa problemática se dá na transformação de textos em vetores ou códigos binários passíveis de serem lidos pela **Inteligência Artificial**.

Para essa transformação e vetorização, recorrentemente é preciso fazer o que, em linguística computacional, denominamos como **pré-processamento**: o preparo de um texto para que ele possa ser analisado e interpretado por um programa. Esse tipo de atividade engloba diversos procedimentos, dentre os quais podemos destacar: remoção de stopwords, stemmização (stemming) e lemmatização.

De forma resumida, a remoção de stopwords trata de retirar do texto palavras que não agregam muito na hora da interpretação, como “o”, “de”, “em”. Já stemmização e lemmatização são procedimentos que, buscando formas básicas e raízes das palavras, facilitam que palavras com sentidos similares sejam entendidas pelo algoritmo.

A partir dessas informações e contextualizações, vocês devem aplicar e estudar como ocorrem tais procedimentos em um *corpus* de âmbito jornalístico. Para tanto, é necessário formular perguntas de análise, como: por que pode existir maior incidência de alguma classe morfológica em uma reportagem política? E determinada stopwords? Como são essas stopwords? Existe algum padrão linguístico sobre sua estrutura? Tais fatos são relacionados ao viés político do meio em que o texto está inserido?

Dessa maneira, além de remover as stopwords do texto e fazer sua stemmização ou lemmatização, vocês devem se valer dessa metodologia de pré-processamento para investigar fatores linguísticos e estruturais do *corpus* definido, refletindo sobre questões tais quais: o que a remoção das stopwords assinaladas pode indicar sobre o *corpus* escolhido? Seus usos geram alguma ambiguidade? Ou faz com que as expressões percam o sentido?

Em suma, vocês precisam investigar a seguinte questão: **O que os procedimentos de linguística computacional nos revelam sobre como textos políticos se diferem? É possível relacionar essas informações com os campos ideológicos que estes textos ocupam?** Os passos, então, são:

- e. Definir dois *corpora*, de orientações políticas diferentes, com no máximo 625 caracteres cada;
- f. Cogitar perguntas e hipóteses sobre parâmetros e características do texto, em suas questões linguísticas e formais;
- g. Aplicar manualmente os procedimentos de remoção de stopwords e lemmatização ou stemmização nas linhas do *corpus*;
- h. A partir do *corpus* pré-processado com os procedimentos aplicados, promover análises e justificar quais as stopwords retiradas e seus motivos, além de interpretar e argumentar a relevância da lemmatização/stemmização feita, condizendo com o *corpus* selecionado e comparando com as hipóteses e perguntas formuladas.
- i. Quais características linguísticas, extraídas do texto pré-processado, podem trazer reflexões sobre o posicionamento temático, contextual e político dos textos selecionados?

Referências para consulta

Campo Olfa Youssef

Olfa Youssef (1964-) é uma pesquisadora de Sousse, Tunísia. Ela se dedica à análise de textos a partir da linguística e da psicanálise, com foco na polissemia dos elementos fundamentais do texto do Corão, no lugar da mulher dentro do Islã e em outros temas envolvendo liberdade religiosa e diálogos inter-religiosos.

FRANÇA, A.; LEMLE, M.; PEDERNEIRA, I.; GOMES, J. **Conexões conceituais: um estudo psicolinguístico de priming encoberto.** *Revista Linguística*, 1(2), 2015. Disponível em: <https://doi.org/10.31513/linguistica.2005.v1n2a4377>

LEITÃO, M. **Psicolinguística Experimental: Focalizando o processamento da linguagem.** In: MARTELOTTA, M. (org.) *Manual de Linguística*. São Paulo: Contexto, 2008. Disponível para Download: <http://www.cchla.ufpb.br/laprol/wp-content/uploads/2014/11/Leit%C3%A3o-2008.pdf>

MAIA, Marcus; LEMLE, Miriam; FRANCA, Anieli Improta. **Efeito stroop e rastreamento ocular no processamento de palavras.** *Ciênc. cogn.* [online], 2007, vol.12, pp. 02-17. ISSN 1806-5821. Disponível em: http://pepsic.bvsalud.org/scielo.php?script=sci_arttext&pid=S1806-58212007000300002&lng=pt&tlng=pt.

MAIA, Marcus. (Org.). **Psicolinguística, psicolinguísticas: uma introdução.** São Paulo: Editora Contexto, 2015.

Campo Nadim

Ibn al-Nadīm (~932-995 EC) foi um erudito de Bagdá, atual Iraque. Ele foi responsável pela compilação do Kitāb al-Fihrist (Catálogo de Livros), um dos mais importantes compêndios medievais, referenciando cerca de 10 mil outras obras e mais de 2 mil autores.

CGEE - CENTRO DE GESTÃO E ESTUDOS ESTRATÉGICOS. **Percepção pública da C&T no Brasil – 2019.** Resumo executivo. Brasília, DF: 2019. 24p. Disponível em: <https://www.cgee.org.br/web/percepcao/faca-sua-analise>.

CATALDI, Cristiane. **A definição utilizada como estratégia divulgativa sobre transgênico na mídia impressa.** *Vertentes*, São João del-Rei, v. 1, n. 32, 2008. Disponível em: https://www.ufsj.edu.br/portal2-repositorio/File/vertentes/Vertentes_32/cristiane_cataldi.pdf

NOCCIOLI, Carlos Alexandre Molina. **O corpo da mulher sob divulgação científica: o senso comum e a reprodução de seus discursos pela mídia.** *PERcursos Linguísticos*, v. 10, n. 25, p. 300-316, 2020. Disponível em: <https://www.periodicos.ufes.br/percursos/article/download/30408/21795/98305>

Campo Fihri

Fatima al-Fihriyya (~800-880 EC) foi uma importante mulher de Kairouan, Tunísia. Ela construiu, na cidade de Fez, Marrocos, a Mesquita de al-Qarawiyyin, depois Universidade de al-Qarawiyyin – considerada a universidade mais antiga em funcionamento e creditada como o centro que introduziu na Europa o uso do zero e dos numerais indo-arábicos.

ENKIN, E.; FORSTER, K. **The Maze Task: Examining the Training Effect of Using a Psycholinguistic Experimental Technique for Second Language Learning.** *Journal of Linguistics and Language Teaching*. v. 5, n. 2, p. 161-180, 2014.

FORSTER, K.; GUERRERA, C.; ELLIOT, L. **The Maze Task: Measuring Forced Incremental Sentence Processing Time.** *Behavior Research Methods*, v. 41, n. 1, p. 163-17, 2009.

FORSTER, K. **Using a maze task to track lexical and sentence processing.** *The Mental Lexicon*, v. 5, p. 347-357, 2010.

LEITÃO, M. M. **Psicolinguística Experimental: focalizando o processamento da linguagem.** In: Mario Eduardo Martelotta. (Org.). *Manual de Linguística*, 1 ed. São Paulo: Editora Contexto, 2008, v. I, p. 217-234.

OLIVEIRA, C.; COSTA, E.; CANABRAVA, K.; BARROS, N. **Examining the use of an online version of the maze task as a pedagogical tool for second language learning.** *Veredas*, v. 24, n. 1, 2020.

OLIVEIRA, C. S. F.; SÁ, T. M. M. **Métodos off-line em psicolinguística: julgamento de aceitabilidade.** *ReVeLe - Revista Virtual dos Estudantes de Letras*, v. 5, p. 77, 2013.

Campo Fakhr-un-Nisa

Fakhr-un-Nisa (1022-1112) foi uma erudita de Bagdá, atual Iraque. Ela foi uma das maiores calígrafas e muhaditas do seu tempo, dedicando-se ao estudo pormenorizado do discurso do profeta Muhammad em uma época em que era bastante incomum mulheres ocuparem esse lugar.

AMOSSY, Ruth (org.). **Imagens de si no discurso: a construção do ethos.** Tradução Dilson F. da Cruz; Fabiana Komesu e Sírio Possenti. São Paulo: Contexto, 2005. Disponível em: https://www.dropbox.com/s/e7sjqmuaz3fuvpeh/imagens_de_si_no_discurso_leia_um_trecho.pdf?dl=0

MAINGUENEAU, Dominique. **Cenas da Enunciação.** Organização: Maria Cecília Péres de Souza-e-Silva e Sírio Possenti. São Paulo: Parábola, 2008.

MAINGUENEAU, Dominique. **Doze conceitos em Análise do Discurso.** Organização: Sírio Possenti e Maria Cecília Péres de Souza-e-Silva. São Paulo: Parábola, 2010.

MAINGUENEAU, Dominique. **Análise de textos de comunicação.** Tradução: Maria Cecília Péres de Souza-e-Silva e Décio Rocha. São Paulo: Cortez, 2013.

MOTTA, Ana Raquel; SALGADO, Luciana (orgs.). **Ethos Discursivo.** São Paulo: Contexto, 2008. Disponível em: https://www.dropbox.com/s/f68hht1jbi84a7r/ethos_discursivo_leia_um_trecho.pdf?dl=0

Campo Jazari

Ismail Al-Jazari (1136-1206) foi um polímata de Jazira (Cizre), atual Turquia. Considerado um “Leonardo da Vinci antes do Leonardo da Vinci”, projetou e inventou dezenas de mecanismos, máquinas e autômatos – o que o rendeu também a alcunha de pai da robótica e da engenharia moderna. Era também um talentoso pintor de iluminuras em manuscritos.

MOTA, Nahendi Almeida; NUNES, Letícia Freitas; MACHADO VIEIRA, Marcia dos Santos. **Você vai ficar roxo de surpresa ao descobrir como intensificamos horrores.** *Roseta*, 2021. Disponível em: <http://www.roseta.org.br/pt/2021/01/29/voce-vai-ficar-roxo-de-surpresa-ao-descobrir-como-intensificamos-horrores/>

SCALDELAI-SALLES, A. L.; SOUZA, E. R. F. de. **Um estudo construcional da microconstrução intensificadora “[x] pra caramba]” no português brasileiro.** *Filologia e Linguística Portuguesa*, v. 22, n. 1, p. 55-79, 2020. Disponível em: <https://www.revistas.usp.br/flp/article/view/163977>

VIEIRA, Sílvia Rodrigues; MACHADO VIEIRA, Marcia dos Santos. **A expressão de grau: para além da morfologia.** *Cadernos de Letras da UFF*, v. 1, p. 63-83, 2008.

Campo Baqir

Taha Baqir (1912-1984) foi um linguista e arqueólogo da Babilônia, Iraque. Ele foi o responsável por decifrar uma série de tábuas matemáticas babilônicas, por traduzir a Epopeia de Gilgamesh do acádio para o árabe e pelas escavações da cidade suméria de Shaduppum, na periferia da atual Bagdá.

ALVES, Nascimento Igor. **Lemmatization vs. Stemming: quando usar cada uma?.** *Alura*. Disponível em: <https://www.alura.com.br/artigos/lemmatization-vs-stemming-quando-usar-cada-uma>

BARION, Eliana Cristina Nogueira. **Mineração de Textos.** *Revista de Ciências Exatas e Tecnologia*. São Paulo, n. 3, v. 3, p. 123 - 140, dez. 2008. Disponível em: <https://www.revista.pgskroton.com/index.php/rcext/article/view/2372>

CAMARELI, Isabella; COSTA, Renato; FROZI, Daniel; FERNANDES, Rian; RODRIGUES, Vitoria. **Conceitos de Linguística para o Processamento de Linguagem Natural - Turing Talks.** Medium. Disponível em: <https://medium.com/turing-talks/conceitos-de-lingu%C3%ADstica-para-processamento-de-linguagem-natural-78241ca2608c>.

GUPTA, Vishal, et. al. **Preprocessing Techniques for Text Mining - An Overview.** *International Journal of Computer Science & Communication Networks*, India, n. 1, v. 5, p. 7 - 16, fev. 2015.

LEMMATIZAÇÃO. *String Fixer*. Disponível em: <https://stringfixer.com/pt/Lemmatization>

STEMMING. *String Fixer*. Disponível em: <https://stringfixer.com/pt/Stemming>