

PRODUÇÃO DE MATERIAL DIDÁTICO COM BASE EM CORPORA:

Uma análise da Linguística de Corpus aplicada ao processo de ensino-aprendizagem de professores de línguas estrangeiras em formação inicial

Jian Nascimento dos Santos (jiannascimento@academico.ufs.br)

Graduando/Universidade Federal de Sergipe

Milena Vitória Leite da Fonseca (milenavitoria@academico.ufs.br)

Graduanda/Universidade Federal de Sergipe

Juliana Pereira Souto Barreto (jpsbarreto@academico.ufs.br)

Doutora/Universidade Federal de Sergipe

Resumo: O principal escopo da Linguística de Corpus (LC) é a língua natural como sistema probabilístico expressa por meio de uma abordagem empírica. Isto é, a visão da linguagem enquanto sistema probabilístico parte do pressuposto de que, apesar de muitos traços linguísticos serem possíveis na teoria eles não ocorrem com frequência. O presente estudo tem como objetivo discorrer sobre a relevância da Linguística de Corpus para o processo de formação docente, seja este inicial ou continuado. Para tanto, analisa-se a aplicação de LC quanto a sua eficiência no processo ativo de ensino-aprendizagem de professores em formação de línguas estrangeiras. Trata-se de uma pesquisa quantitativa com exploração em ferramentas de corpora como o COCA (The Corpus Of American English) e AntConc, com posterior análise qualitativa dos dados. Tomar-se-á por base leituras relativas à Linguística de Corpus como Sardinha (2004), Cantarotti (2017), Delfino (2016) dentre outros. Para o desenvolvimento desta pesquisa, um questionário foi aplicado com professores em formação em que, com base nos resultados obtidos, deu-se o processo de criação de atividades com suporte das plataformas aqui escritas. Como resultado, observa-se um campo ainda pouco explorado, mas com grandes possibilidades de desenvolvimento no que se refere a formação de professor de línguas estrangeiras, principalmente, quando voltada à produção de material didático.

Palavras-chave: Linguística de corpus; material didático; línguas estrangeiras.

Introdução

A priori, este estudo foi iniciado com o propósito de discorrer sobre a relevância do estudo da Linguística de Corpus no processo de formação inicial de dois calouros no curso de Letras Inglês na Universidade Federal de Sergipe, que se deu por meio do seguinte plano de trabalho (projeto de extensão): observação, análise e produção de material de/para ensino-aprendizagem de línguas/literaturas estrangeiras com base em corpora (ALEC). Portanto, muito do que está aqui discorrido provém das práticas deste projeto de extensão, que

ocorreu de forma remota durante o período pandêmico, entre 2021 e 2022. O desenvolvimento do projeto de extensão, a pesquisa e posterior análise das ações elaboradas são desenvolvidas a partir do envolvimento de alguns estudantes, dois deles autores desse texto e participantes do projeto, na época, ainda calouros, em períodos iniciais do curso.

Apesar de a Linguística de Corpus (doravante LC) ser um campo ainda muito recente no Brasil, temos pesquisadores particularmente interessados no tema, desenvolvendo trabalhos, e cujos resultados já contam com relevante destaque para pesquisa internacional. Assim, quando se trata do uso de ferramentas de manuseio com corpora não há como negar a necessidade de um referencial teórico internacional devido a sua criação e maior desenvolvimento fora do país. Não há como falar das ferramentas de corpora, ou seja, de LC, sem antes falar um pouco mais sobre o que se trata a própria Linguística de Corpus. Tony Berber Sardinha (2004), um dos pioneiros da LC no Brasil, define a Linguística de Corpus como Linguística Empírica visto que se trata da observação da linguagem reunida em uma coletânea de textos - corpus - e da análise do uso da língua como sistema probabilístico, em consonância com Halliday (1993), e não como um sistema de possibilidades, tendo assim uma certa dissonância com a Linguística Chomskyana.

Como exposto por Berber Sardinha (2004, p. 30), “A visão da linguagem como sistema probabilístico pressupõe que, embora muitos traços linguísticos sejam possíveis teoricamente, não ocorrem com a mesma frequência”. Ou seja, não é porque certas estruturas são possíveis de acontecer no vernáculo que estas serão as mais recorrentes, logo, as mais usadas pelos falantes da língua em questão. Em um primeiro olhar, é compreensível a confusão entre probabilidade com possibilidade, sendo justamente nesse ponto em que a Linguística de Corpus diverge: de um lado temos Halliday com a probabilidade e Chomsky com a possibilidade. No entanto, em sua obra "Linguística de Corpus", Berber Sardinha (2004) deixa claro a diferença e esclarece que a LC trabalha conforme a visão de Halliday uma vez que se trata da visão probabilística da língua em ação. Desse modo, é perceptível o potencial dessa área de pesquisa no ensino-aprendizagem de idiomas, uma vez que proporciona uma compreensão minuciosa das particularidades de uma língua, a exemplo da observação de concordâncias, colocações e padrões linguísticos presentes no corpus analisado.

Sobre essa ótica, no que se refere a atuação de professores, a Linguística de Corpus vem sendo cada vez mais usada para a preparação de materiais didáticos, como é visto em trabalhos de pesquisadores como Berber Sardinha; Delfino e Rampaso (2017), Delfino (2016) e Paulino (2016). Isso porque muitas das atividades e exercícios disponíveis nas escolas, onde há ensino de línguas, seguem métodos de ensino tradicionalistas que tendem a não gerar o engajamento dos alunos. Além disso, Paulino (2016 p. 93) argumenta que:

Não obstante à falta de estrutura e às salas cheias, os materiais didáticos, quando disponibilizados para o trabalho, geralmente não atendem às necessidades do público alvo, o que demanda do professor o trabalho de produzir vários exercícios para o trabalho em sala de aula, os quais se baseiam, de forma geral, nas consultas à diferentes livros didáticos ou em buscas aleatórias na internet.

Metodologia

Para a realização desse trabalho, partimos inicialmente da análise dos dados obtidos por meio de um questionário eletrônico, o qual foi compartilhado com a comunidade acadêmica da Universidade Federal de Sergipe. O questionário, que continha perguntas relacionadas à percepção de professores e graduandos atuantes em sala de aula sobre aspectos pertinentes ao ensino de línguas estrangeiras, tinha por objetivo levantar informações sobre as necessidades desses professores em formação inicial quanto a material a ser utilizado em sala de aula. Só depois de coletadas as informações, demos início a produção de materiais didáticos, a qual foi realizada utilizando não apenas as demandas observadas no questionário como também fazendo uso, quando necessário, de ferramentas como COCA e AntConc com o intuito de aplicar na prática os princípios da Linguística de Corpus e, desse modo, analisar sua eficiência no ensino de línguas.

Análise Dos Dados Obtidos

O questionário utilizado para esta pesquisa em específico foi aplicado exclusivamente com estudantes de letras ainda em processo de formação, todos esses vinculados ao departamento de letras estrangeiras. Vale salientar que muitos deles participavam de algum programa de extensão na universidade, assim, pode-se observar no gráfico abaixo:

Gráfico 1: programas em que os professores estavam inseridos.

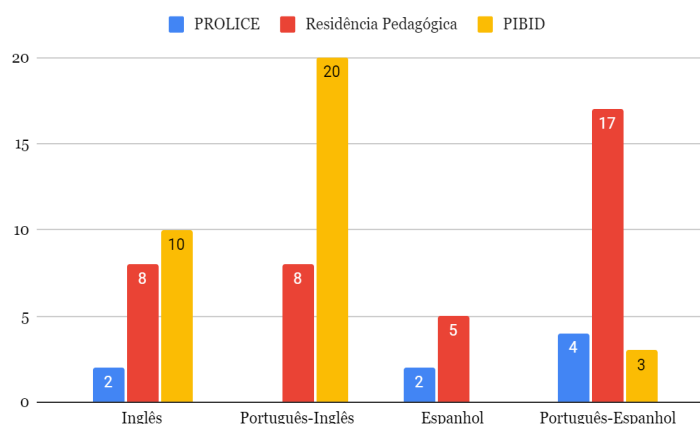


Gráfico do formulário aplicado para professores em formação na pesquisa supracitada.

A distinção destes programas não pareceu ter relevância para esta pesquisa além do fato de que os professores em formação estavam experienciando a sala de aula em primeira mão, ou seja, estavam em contato com a prática. As perguntas eram direcionadas ao que os professores julgavam ser necessário e desejado pelos alunos em sala de aula. Como por exemplo, o gênero com o qual os alunos poderiam se envolver mais no processo de ensino-aprendizagem da língua alvo.

Gráfico 2: gêneros de interesse dos alunos.

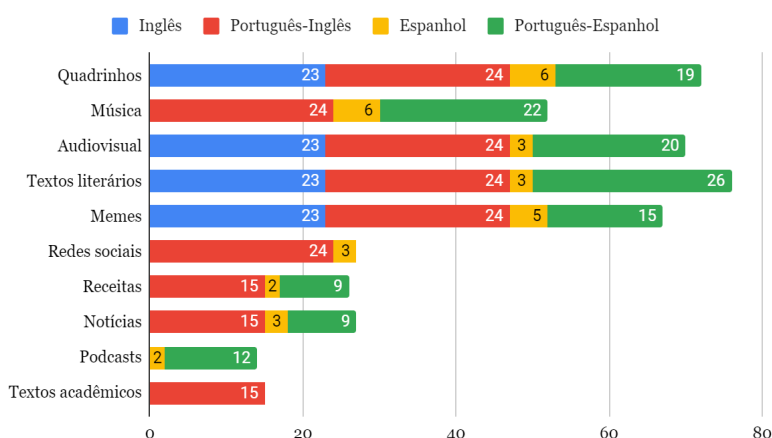


Gráfico do formulário aplicado para professores em formação no plano de trabalho supracitado.

Esse tipo de informação foi primordial para o processo inicial de desenvolvimento das atividades, pois foi a partir destas informações que pudemos nos basear sobre o que poderia tomar como foco principal para o desenvolvimento das nossas atividades. Porém, não foi deixado de lado aspectos como o nível escolar, idade, se o professor pretendia focar em oralidade ou escrita e linguística ou literatura, recursos disponíveis ao professor, recursos disponíveis aos alunos, objetivo para o desenvolvimento da língua, temas a serem abordados, etc. Foi possível notar uma certa discrepância quanto ao foco do professor, como é possível observar nos gráficos abaixo:

Gráfico 3

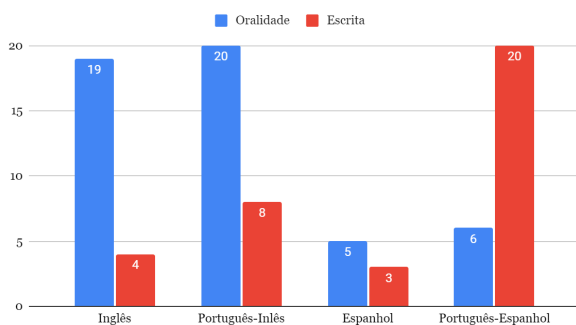
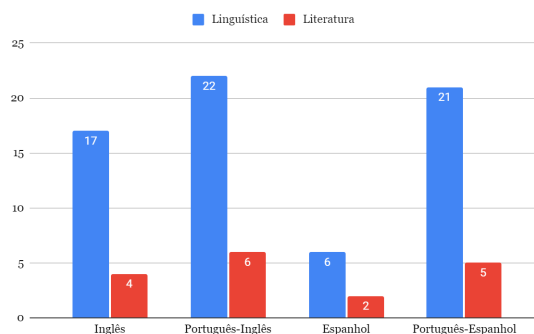


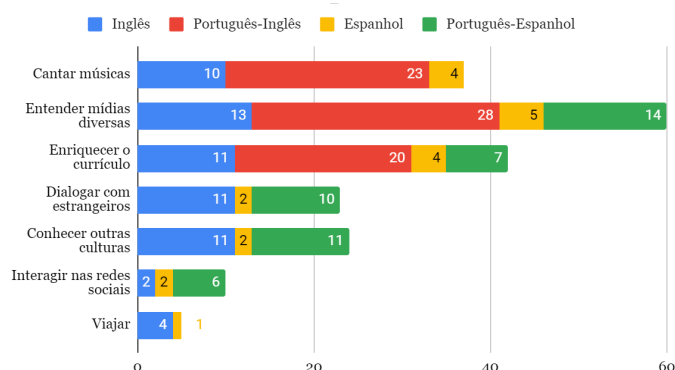
Gráfico 4



(Dados obtidos do formulário aplicado para professores em formação na pesquisa supracitada.)

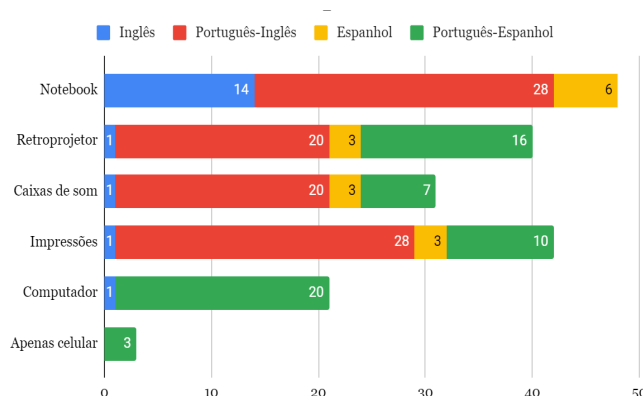
Esses dados foram importantes para afunilar ainda mais a criação de materiais, ou seja, se as atividades seriam focadas em oralidade ou escrita e se o foco seria trabalhar mais língua ou literatura, pelos questionários foi possível identificarmos as necessidades dos alunos. Entender o propósito dos alunos no processo de aprendizagem também foi enriquecedor no levantamento de dados.

Gráfico 5: Motivação dos alunos.



Quando se entende o propósito fica muito mais fácil delimitar os meios para se chegar no objetivo desejado. Além disso, foi primordial saber como as escolas eram descritas em relação aos seus recursos, ou seja, a visualização dos recursos disponíveis para os professores, visto que esse é um fator de suma importância, que deve ser considerado quando na posição de editor, planejador elaborador e/ou criador de qualquer tipo de atividade.

Gráfico 6: recursos disponíveis para os professores.



Todavia, ainda que considerando todos esses aspectos, esse questionário foi usado apenas como ponto de partida para a criação das atividades, as quais foram estruturadas com base em corpora enquanto explorando de forma mais objetiva as ferramentas de corpora.

Sem desconsiderar todos os levantamentos feitos até esse ponto, somadas as base em corpora que tomaríamos, nós ainda tomaríamos uma última decisão que foi a de atentar para o gênero textual, uma vez que acreditamos ser um dos principais pontos a levar o aluno à construção de sentido de forma autônoma. Ao apresentar a língua de forma situada ao aluno, ou seja, inserida em um gênero textual, a língua deixaria de ser um conteúdo a ser tomado e passa a ser um processo, ou seja, a partir do reconhecimento e identificação do gênero, o aluno recorre aos seus conhecimentos de como se comportar linguisticamente naquela situação, o que já o ajuda imensamente, tendo apenas que desenvolver, a partir daí, uma habilidade de comunicação dentro de uma situação e, agora, com um propósito.

Assim, optamos por fazer uso de gêneros discursivos por acreditarmos que, para além de todo o contexto dado, os alunos também se sentem mais instigados a aprender, sua motivação e aspectos como os recursos disponíveis devem ser considerados - não apenas na LC, mas em qualquer tipo de processo de ensino e aprendizagem.

O Material Didático Produzido

Dentre os materiais didáticos elaborados, a atividade "*Whodunnit?*" [Apêndice 1] foi escolhida para ser exposta neste artigo. O título que remete diretamente ao subgênero de romance policiais, em que se busca encontrar o culpado de um crime, não apenas busca indicar a temática de detetives da aula como também já apresenta o tópico gramatical a ser trabalhado, o uso de palavras interrogativas. A criação desse material didático partiu inicialmente da intenção de trabalhar com *question words* e a relação com a literatura detetivesca surgiu com uma possibilidade de contextualizar o conteúdo, gerando maior engajamento dos alunos, além do uso da língua em um contexto específico.

Com isso em mente, buscamos no YouTube vídeos sobre detetives, tentando encontrar alguma comparação entre os dois mais famosos detetives da literatura mundial, Sherlock Holmes de Sir. Arthur Conan Doyle e Hercule Poirot de Agatha Christie, os quais julgamos ser uma excelente introdução para a temática. Assim que encontramos um vídeo adequado, aqui referido como Vídeo 2 (KAH YEE, 2017), submetemos a transcrição de seu áudio no *software* livre *Corpus of Contemporary American English* – COCA para a análise das palavras mais frequentes no texto. A coocorrência de *question words* foi menor do que o esperado, ainda assim notamos o grande potencial do vídeo para a elaboração de perguntas usando palavras interrogativas já que o texto apresentava muitas informações sobre Holmes e Poirot.

Desse modo, selecionamos dois outros vídeos de curta duração para fornecer mais exemplos de uso de *question words* e promover a ampliação do vocabulário dos alunos. Enquanto um dos vídeos, referido como Vídeo 1 (LINKYTHINK, 2018), tratava de perguntas comuns feitas por detetives, o outro era um trecho do filme "*Batman: The Dark Knight* (2008)", referido como Vídeo 3 (FLASHBACK FM, 2019), no qual ocorre o interrogatório do vilão Coringa pelo super-herói Batman. Diante de tantos exemplos diferentes, decidimos selecionar boa parte deles para a criação de uma lista de palavras que poderia ser usada pelos alunos de diversas formas, desde uma simples consulta ou uma análise minuciosa da estrutura das frases, observando os padrões que se repetem com frequência na lista.

Em sua versão final, a atividade começa despertando o conhecimento prévio acerca da temática detetive. Para isso, utilizamos algumas imagens e questionamentos relacionados ao mundo das investigações policiais com a finalidade de não apenas preparar os alunos para as

ações que seriam solicitadas mais a frente como também já seriam destacadas as *question words* por meio dessas perguntas, as quais estão escritas em inglês e contemplam diferentes exemplos de palavras interrogativas.

Em seguida, sugerimos a exibição do Vídeo 1 para agregar mais informações sobre detetives e estimular a *listening* da turma com mais exemplos de *question words*. Após essa apresentação inicial, é chegada a hora de convidar os alunos a se tornarem protagonistas, ou seja, a se tornarem os próprios detetives em busca de informações sobre os detetives Sherlock Holmes e Hercule Poirot, os quais estarão disponíveis no vídeo seguinte a ser exibido.

Na atividade, sugerimos uma divisão da turma em grupos, e, então, que cada grupo assuma um detetive para buscar informações e elaborar perguntas em inglês com *question words* para o outro grupo responder, o qual também deve fazer o mesmo com as informações do seu detetive. Devido às diferentes realidades possíveis em uma sala de aula, optamos por deixar que os professores que fizessem uso desse material didático definissem como ocorreria a divisão da turma. Além disso, disponibilizamos a transcrição do Vídeo 2 para substituir a exibição do vídeo caso fosse preciso, ou até mesmo fosse usada em conjunto para facilitar a busca pelas informações. Vale destacar que o Vídeo 3 acabou servindo mais com um complemento opcional que pode não ser exibido se o professor julgar não ser necessário ou se o tempo da aula não for o suficiente.

Conclusões

Dada a experiência que esta pesquisa nos proporcionou, observamos a relevância da utilização da LC e da utilização das ferramentas de manuseio de corpora para a construção de atividades voltadas ao ensino aprendizagem de línguas estrangeiras. Observamos também, certa limitação de conhecimento e de possibilidades de produção de conhecimento na área. Tendo apresentado o trabalho em espaços de compartilhamento de pesquisas acadêmicas, como o VII Sefeli, observamos pouco ou nenhum desenvolvimento de pesquisas na área em nossa região. Uma das respostas possíveis para esse escasso desenvolvimento esteja no fato de que a LC dialoga com campos bem diversos, a exemplo da Ciência de Computação, que realiza o processamento da linguagem natural. Todavia, apesar do campo ainda pouco explorado, observamos grandes possibilidades de desenvolvimento no que se refere a formação de professor de línguas estrangeiras, principalmente, quando voltada à produção de material didático.

Foi possível concluir que antes de qualquer inferência intuitiva sobre a língua, vale a pena realizar testes de checagem a fim de verificar se, de fato, o que acreditávamos aparecer no texto de fato aparece, a exemplo do que o que nos aconteceu quando da construção da atividade aqui relatada. Ou seja, para a construção de atividades cujos conteúdo nós acreditamos já saber quais são, sempre vale a pena conferir por meio de ferramentas de corpora se nossa intuição de fato corresponde a realidade. Concluimos portanto que a LC nos possibilita uma visualização da língua que direciona o nosso olhar para uma análise de forma mais autônoma, que permite e promove um olhar de pesquisador sobre a língua e não apenas de reprodutor desta.

Devemos lembrar que, assim como tudo evolui, as metodologias e abordagens também devem evoluir. Entretanto, apesar dos estudos evoluírem, não significa que os recursos e as singularidades evoluem em conjunto. Concluimos, por fim, que a linguagem natural, bem como os gêneros discursivos, auxiliados pela LC, podem ser usados para produção de material didático no ensino e aprendizado de línguas estrangeiras para o desenvolvimento de atividades mais contextualizadas e menos fragmentadas. Esses podem servir de base para uma pedagogia que destaca a língua a serviço da comunicação, ou seja, como uma forma de comunicação.

REFERÊNCIAS

BERBER SARDINHA, Tony. **Linguística de corpus**. São Paulo: Manole, 2004

BERBER SARDINHA, T.; DELFINO, Maria Claudia; RAMPASO, Marianne. **Preparação de Material Didático para o Ensino de Línguas com Base em Corpora**. São Paulo: The Specialist, 2017.

BIBER, D. , 1988. **Variation across speech and writing**. Cambridge: Cambridge University Press _____, 1995. **Dimensions of Register Variation - A Cross-Linguistic Comparison**. Cambridge: Cambridge University Press.

CANTAROTTI, ALINE. **Data-driven learning, tradução e secretariado executivo: proposta de abordagem de auto-estudo para aprendizes de inglês**. REVISTA ENTRELÍNGUAS , v. 3, p. 227-247, 2017.

DELFINO, M. C. N. , 2016. **Uso de Música para o Ensino de Inglês como Língua Estrangeira em um Ambiente Baseado em Corpus**. (Dissertação de Mestrado) - LAEL, PUCSP.

FLASHBACK FM. **Batman interroga Coringa | O Cavaleiro das Trevas**. YouTube, 12/11/2019. Disponível em: https://youtu.be/jane6C4rIwc?si=9e6Pp_iACpPUIwSi.

HALLIDAY, M. A. K. 1993. **Quantitative studies and probabilities in grammar**. In M. Hoey (Org.), *Data Description Discourse -- Papers on the English Language in Honour of John McH Sinclair on his Sixtieth Birthday*. London: HarperCollins, p. 1-25.

PAULINO, Eliene de Souza. Produção de atividades didáticas com o uso de corpora: uma experiência com professores de língua inglesa de escolas públicas. p. 93-97. In: **Anais do EBRALC 2015 & ELC 2015 [Blucher Social Science Proceedings, n.3 v.2]**. São Paulo: Blucher, 2016.

TED-ED. **Who is Sherlock Holmes - Neil McCaw**. YouTube, 05/05/2016. Disponível em: https://youtu.be/I8992A5oAWM?si=-IIVEqq6H_uZ9tok.

APÊNDICE 1

Activity: **Whodunnit?**

- Look at the pictures:



- What** do you know about these pictures?

- Why** do you think people commit crimes?

- What** do you think about detectives? **Where** did you hear about them?

- Do you know **when** detective stories came about?

- In crime novels, **what** do you think this expression "*whodunnit*" means?

- What kind of** questions do detectives ask? Watch the video:

[Detective Questions: Comprehending Inference](#)

Now it's your turn to become a detective!!!

- Ask questions;
- Gather all information;
- Find the clues.

New Case : Sherlock Holmes Vs Hercule Poirot!
Get ready for the INTERROGATORY!

You need to **gather as much information** as possible about one of these famous detectives and **create questions from them**. Read the text of the next page to find out more clues about Holmes and Poirot.

After that, question your colleagues with these questions.

But first, ...

- **How much** do you understand about interrogations? Watch this video:

[Batman interrogates the Joker | The Dark Knight \[4k, HDR\]](#)

- **How** safe are you to ask your own questions?

If you need, you can use this list of expressions:

Who did it? - **Quem** fez isso?

When did they do it? - **Quando** eles fizeram isso?

Where did they do it? - **Onde** eles fizeram isso?

Why do you think Bob rolled his eyes on line 27? - **Por quê** você acha que Bob revirou os olhos na linha 27?

What kind of animal **do you** think Hilda **is** and **why**? - **Que tipo de** animal você acha que Hilda é e por quê?

Where is Bob from? - De onde Bob é?

Whose is this lost animal? - **De quem** é o animal perdido?

Which of the lines **did** Bob roll his eyes? - **Quais** das linhas Bob revirou os olhos?

How was Hilda found? - **Como** Hilda foi encontrada?

What have **you** done with him? - **O que** você fez com ele?

Who did you leave him with? Your people? - **Com quem** você o deixou? (**Com**) seus capangas?

Does it depress you, Commissioner? To know just **how** alone **you** really **are**? - Isso o deprime, Comissário? Saber **o quanto** sozinho você realmente está?

Where is he? - **Onde** ele está?

What 's the time? - **Que** horas são?

What the difference **does** that make? - **Que** diferença isso faz?

Then **why do you** want to kill me? - Então, **por quê** você quer me matar?

What would I do without you? Go back to ripping off mob dealers? - O que eu faria sem você? Voltar a roubar os mafiosos?

Who 's in control? - **Quem** está no controle?

Which is ? - **Qual** é?

Where are they ? - **Onde** eles estão?

Which one **are you** gone after? - **Qual** deles você foi atrás?

Sherlock Holmes Vs Hercule Poirot

If you're a fan of crimes and mysteries, it's no doubt that you'll recognize these names. Today, we are going to talk about the world-famous detectives, Sherlock Holmes and his most comparable detective, Hercule Poirot. Holmes and Poirot are fictional characters written by Sir (Arthur) Conan Doyle and Agatha Christie during the 19th century. Holmes and Poirot are both private detective bases in London, the capital city of England. Holmes was active during the Victorian era and Poirot was active during the Edwardian era. Holmes is a pale looking, sharply-featured British man. He is over 6 feet tall and often depicted in a brown cap, deerstalker hat and his curved pipe. On the other hand, Hercule Poirot is an elderly Belgian man with a wax mustache. He is only 5 feet 4. Unlike Holmes, Poirot often dressed in a silk suit with a bow tie and a hat. Both this detective has a sidekick who was a stereotypical British gentleman. While Holmes had Dr. Watson, Poirot enjoyed the company of Captain Arthur Hastings. Besides, Holmes and Poirot have strong feelings for only one woman, both who were strangely on the other side of the law.

While it was Countess Vera Rossakoff for Poirot, Holmes was attached to Irene Adler. However, these relationships did not go far. People say detectives are hardly ever quite sane, this is true about Sherlock Holmes and Hercule Poirot. They both have certain habits and hobbies which seem weird compared to ordinary people. Holmes is obsessed with chemistry, Poirot is obsessed with his mustache and cleanliness. Holmes loves to play the violin, Poirot loves to rearrange his wardrobe. Holmes experimented himself with drugs, Poirot indulged himself in delicious and fine food. Yet, both of them have one thing in common - A passion to unravel mysteries. There is a great difference between Sherlock Holmes and Hercule Poirot when it comes to solving crimes. Holmes is a brilliant man with a deep knowledge of science and chemistry. Nothing in his life is more important than science and the whole world is a certain laboratory for him.

When solving crimes, Holmes uses logical deduction and collects every single physical evidence and examines them. Furthermore, Holmes it's said to be the master of disguise. He disguised to gather evidence undercover so convincingly that even Watson fails to penetrate them. Holmes sees crimes only as a puzzle he has to solve. On the other hand, Hercule Poirot does not have a scientific mind and a definite method. Nevertheless, he's an expert on human behavior. Poirot's deductions are mainly based on his excellent knowledge of human psychology. He always examines the personalities, behaviors and the relationships of the people involved in the case. Poirot is also a clue based detective depending on logic, which is represented in his vocabulary by his famous phrase - The little gray cell. Sherlock Holmes and Hercule Poirot continue to captivate its readers all over the world. Both the characters enjoy unprecedented fame and are no doubt the two finest crime solvers in history.

Video transcript by [Sherlock Holmes VS Hercule Poirot](#)