**🧠 Como o Código Funciona**

**1. Pré-processamento**

Antes de tentar descriptografar o texto, o programa faz:

* **Remove acentos** (ex: "ação" vira "acao").
* **Remove símbolos estranhos** (mantém só letras e espaços).
* **Transforma tudo para minúsculas**.

Função:

preprocess\_text(text)

**2. Tentativa de Decriptar**

O programa **testa todos os shifts possíveis** (de 1 até 25) na cifra de César.

Pra cada tentativa:

* Descriptografa o texto.
* Verifica se o texto **faz sentido**.

Função:

caesar\_decrypt(text, shift)

**3. Validação Semântica**

Depois de descriptografar, ele precisa saber se o texto "tem palavras de verdade".

Aqui ele faz duas checagens:

**(a) Usar a GLC para validar separação de palavras:**

* A **GLC (Gramática Livre de Contexto)** que criamos é **super simples**.
* Ela só diz: **"uma frase válida é uma sequência de palavras separadas por espaços"**.
* Não importa quais palavras, desde que sejam separadas.

Parte da gramática:

start: WORD (WORD)\*

👉 Ou seja, **uma ou mais palavras**.

Se o texto descriptografado **não for palavras separadas**, a GLC dá erro e descartamos esse shift.

**(b) Usar o dicionário português:**

* Comparamos palavra a palavra no texto descriptografado.
* **Contamos** quantas palavras existem no nosso portuguese\_words.txt.
* Escolhemos o shift que **tem mais palavras corretas**.

Função:

count\_valid\_words(text, dictionary)

**4. Seleção Final**

Depois de testar todos os 25 shifts:

* O shift que conseguiu **mais palavras corretas** e passou na GLC é o escolhido.
* Se nenhum shift for bom, avisa que não foi possível descriptografar.