

**UNIVERSIDADE FEDERAL DA FRONTEIRA SUL  
CAMPUS CHAPECÓ  
CIÊNCIA DA COMPUTAÇÃO  
LINGUAGENS FORMAIS E AUTÔMATOS  
PROFESSOR: ANDREI DE ALMEIDA SAMPAIO BRAGA**

**ANALISADOR DE FRASES**

Yuri Luis Malinski Lanzini – 2221100006

## **1. Descrição**

Este analisador sintático, desenvolvido com a ferramenta YACC, tem como objetivo principal validar a estrutura gramatical de frases simples em português. Ele opera em conjunto com um analisador léxico (gerado pelo Lex) que realiza a tokenização da entrada, identificando e classificando os elementos lexicais (artigos, pronomes, verbos, adjetivos, advérbios e substantivos).

### **Funcionamento**

1. **Análise Léxica:** O analisador léxico processa a entrada (arquivo de texto ou entrada padrão), segmentando-a em tokens e classificando cada token de acordo com sua categoria gramatical.
2. **Análise Sintática:** Com base em uma gramática livre de contexto predefinida, o analisador sintático verifica se a sequência de tokens constitui uma frase válida. Ele constrói uma árvore sintática que representa a estrutura da frase, assegurando que cada elemento esteja na posição correta.
3. **Validação:**
  - **Erros Sintáticos:** Se a sequência de tokens não se adequar a nenhuma regra da gramática (ex: um verbo no lugar de um substantivo), o analisador sintático reporta um erro, indicando o token problemático e a linha correspondente.
  - **Erros Léxicos:** Se o analisador léxico encontrar uma palavra não classificável (ausente na tabela de símbolos e não correspondente a nenhuma palavra-chave), ele reporta um erro léxico, indicando a palavra desconhecida.

### **Limitações**

- **Gramática Limitada:** A gramática atual não engloba todas as nuances e construções da língua portuguesa, restringindo a variedade de frases analisáveis.
- **Estruturas Complexas:** O analisador não lida com ambiguidades ou frases com estruturas mais complexas, como orações subordinadas ou coordenadas.
- **Erros Semânticos:** Frases gramaticalmente corretas, mas sem sentido (ex: "A pedra comeu o livro") não são detectadas.

## 2. Gramática livre de contexto

**texto:** Representa o texto completo, que pode ser uma sequência de frases ou uma única frase.

**frase:** Define a estrutura básica de uma frase: sujeito, verbo e objeto.

**sujeito:** Pode ser um pronome ou um artigo seguido de um substantivo.

**objeto:** Pode ser um artigo seguido de um substantivo, um adjetivo ou um advérbio.

Gramática livre de contexto:

1. texto → texto frase
2. texto → frase
3. frase → sujeito VERBO objeto
4. sujeito → PRONOME
5. sujeito → ARTIGO SUBSTANTIVO
6. objeto → ARTIGO SUBSTANTIVO
7. objeto → ADJETIVO
8. objeto → ADVERBIO

## 3. Testes

Os testes foram realizados no terminal do Linux. Para compilar, foi necessário ir pelo terminal para o diretório onde estão todos os arquivos e executar o comando ‘make’, após isso ./frases (nome do arquivo.txt com os dados que deve estar no mesmo diretório).

Inicialmente, verificou-se o arquivo de entrada (input) fornecido pelo próprio autor, confirmando que o analisador funciona conforme esperado.

Dados do arquivo de entrada:

```
1 pron eu tu ele ela nos vos eles elas
2 arti o a os as um uma uns umas
3 subs chuva colega papel cadeira caixa
4 verb cai viu pula tem somos fazem
5 adv aqui ali adiante embaixo perto longe
6 adj velho novo alto baixo forte fraca seco molhado bonito
7
8 a chuva cai forte
9 ele viu uma caixa
10 eu cai longe
11 o colega pula alto
12 eu cai na lama viu
```

As quatro primeiras frases são válidas, enquanto a última, "eu cai na lama viu", apresenta dois erros lexicais e um erro sintático:

```
yuri_lanzini@Linux-Yuri:~/Área de Trabalho/Compiladores/Labs/Lab25/Frases$ ./frases input
A frase é válida.
A frase é válida.
A frase é válida.
A frase é válida.
Erro léxico: "na" não reconhecido
Erro léxico: "lama" não reconhecido
Erro sintático: símbolo "viu" (linha 12)
yuri_lanzini@Linux-Yuri:~/Área de Trabalho/Compiladores/Labs/Lab25/Frases$
```

Foi criado outro arquivo de entrada chamado “testes” para analisar frases gramaticalmente corretas, mas sem sentido:

Dados do arquivo testes:

```
1 pron eu tu ele ela nos vos eles elas
2 arti o a os as um uma uns umas
3 subs menino árvore casa livro carro mesa sofa
4 verb corremos leu pinteí escreveu brinca comeu
5 adv rapido devagar ontem hoje longe perto
6 adj grande pequeno azul verde feliz triste
7
8 eu corremos a casa
9 ele brinca um sofa
10 nos pinteí longe
11 o carro leu um menino
12 o livro comeu a mesa
```

Todas as frases foram validadas mesmo sem sentido, conforme o esperado.

```
yuri_lanzini@Linux-Yuri:~/Área de Trabalho/Compiladores/Labs/Lab25/Frases$ ./frases testes
A frase é válida.
A frase é válida.
A frase é válida.
A frase é válida.
A frase é válida.
yuri_lanzini@Linux-Yuri:~/Área de Trabalho/Compiladores/Labs/Lab25/Frases$
```

Durante os testes, foi identificado outro erro: se ocorrer um erro sintático e tiver frases abaixo, o programa não verifica as frases subsequentes e é finalizado, porém se o erro for léxico o programa funciona como deveria(indicando os erros e verificando as próximas frases).

Arquivo testes (Modificado para gerar o erro):

```
1 pron eu tu ele ela nos vos eles elas
2 arti o a os as um uma uns umas
3 subs menino árvore casa livro carro mesa sofa
4 verb corremos leu pintei escreveu brinca comeu
5 adv rapido devagar ontem hoje longe perto
6 adj grande pequeno azul verde feliz triste
7
8 (gerar erro aqui) leu
9 eu corremos a casa
10 ele brinca um sofa
11 nos pintei longe
12 o carro leu um menino
13 o livro comeu a mesa
```

Erro descrito acima:

```
yuri_lanzini@Linux-Yuri:~/Área de Trabalho/Compiladores/Labs/Lab25/Frases$ ./frases testes
Erro léxico: "gerar" não reconhecido
Erro léxico: "erro" não reconhecido
Erro léxico: "aqui" não reconhecido
Erro sintático: símbolo "leu" (linha 8)
yuri_lanzini@Linux-Yuri:~/Área de Trabalho/Compiladores/Labs/Lab25/Frases$
```

## 4. Referências

<https://github.com/JudsonSS/Compiladores/tree/master/Labs/Lab25/Frases>

<https://www.youtube.com/watch?v=mGRN6WagUKM>