The Role of Dictionary Users in Detecting Neologisms: A Case Study with *Dicionário da Língua Portuguesa (DLP)*

Maria Leonor Reis
Instituto de Lexicologia e Lexicografia da Língua Portuguesa
Academia das Ciências de Lisboa (ACL)
mleonoreis@gmail.com

Introduction

The **digital transition** has mitigated many constraints of printed dictionaries (space limitations and difficulties in updating and reviewing lexical content).

However, institutions committed to lexicographic publication – like ACL. – face significant challenges, including managing vast amounts of linguistic data with limited resources.

The DLP-ACL incorporates user interaction through direct methods (e.g., user suggestions and feedback) and indirect methods (e.g., log file analysis).

Research Question

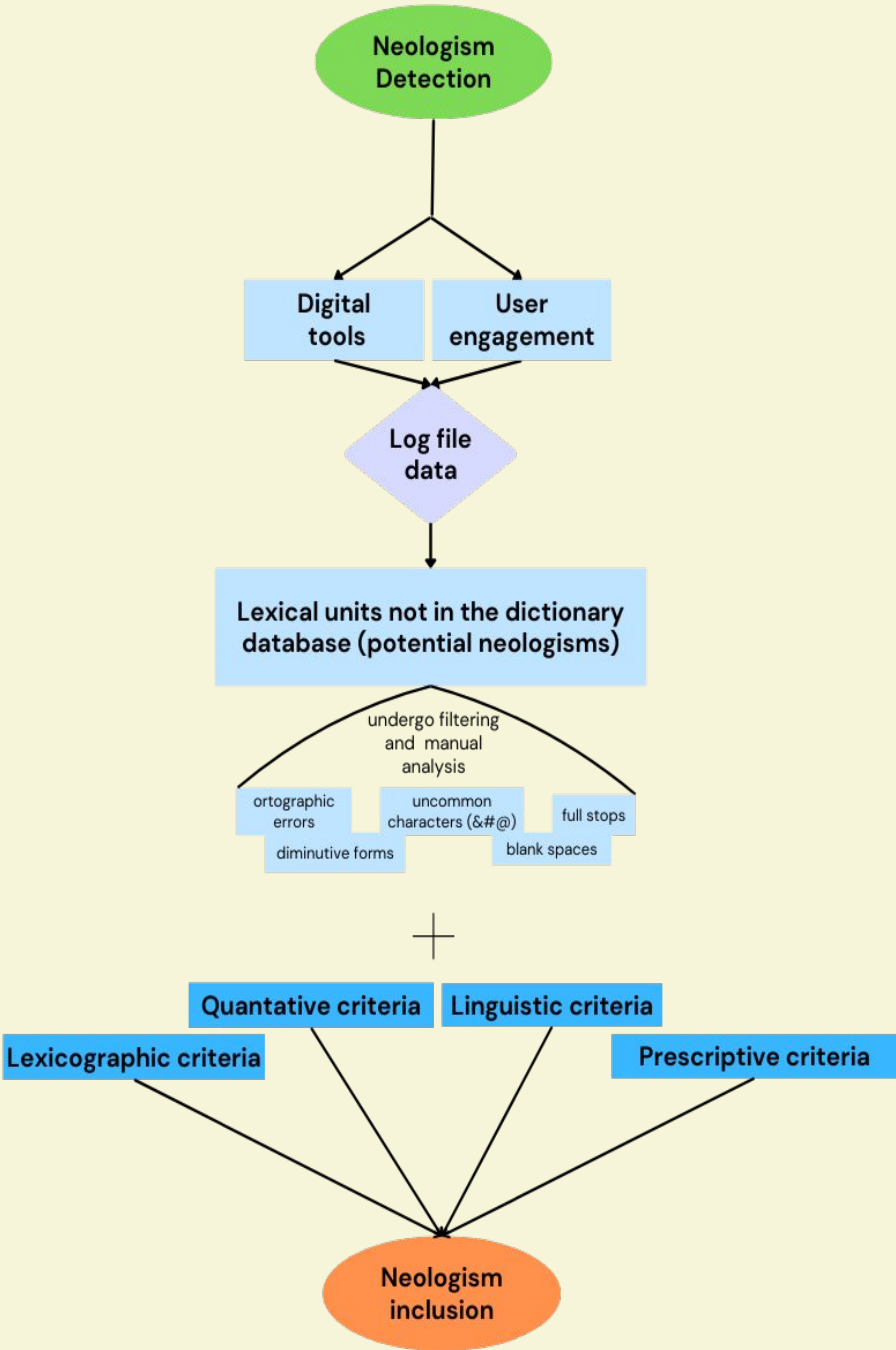How can **user engagement** and **digital resources** facilitate the detection of neologisms?

Resources

1. DLP–ACL
   https://dicionario.acad-ciencias.pt/

1. Log file data – Three levels

- **0**: lexical units not in the dictionary database **(potential neologisms)**
- **1**: lexical units already registered as entries **(potential semantic neologisms)**
- **2**: lexical units present in the dictionary database but requiring lexicographic work

Data analysed

Log data **level 0**

From 01/01/24 to 28/11/24 → **429 024 searched words**

Methodology

Results

– Neologisms: *aerofilme; avódrasta*; hidrospeed; *subsidiodependente; autoabandono; neuroprótese; radiotelefone; biossistema*; biodesign; cardiofitness; fashionable
– Missing entries (gaps): neoliberal; pandã; involucrar; balaclava

Challenges and Future Work

- The dataset is (still) too large for manual analysis.
- Data remains noisy despite filtering, with residual toponyms, anthroponyms, foreign words, plural and flexionate forms within the morphological analyser.
- Developing a comprehensive and up-to-date reference corpus for Portuguese.
- Cross-referencing log file data with the *Vocabulário Ortográfico da Língua Portuguesa* (VOLP) to eliminate anthroponyms and toponyms.

Action
ENEOLI
European Network
on Lexical Innovation

Affiliation