TRATAMENTO DO CORPUS C-ORAL-QUILOMBOLA: TRANSCRIÇÃO, REVISÃO, ALINHAMENTO TEXTO-SOM

AUTOR: PEDRO HENRIQUE MONTEIRO DA SILVA

CO-AUTOR/ORIENTADOR: LUCIANA BEATRIZ BASTOS AVILA

Resumo: O projeto de pesquisa “C-ORAL-Quilombola: compilação de corpus de fala espontânea de comunidades remanescentes de quilombos no Extremo Sul da Bahia” visa à compilação de um corpus de fala espontânea informal do português de uma comunidade remanescente de quilombo, especificamente na comunidade de Volta Miúda, no município de Caravelas, extremo sul da Bahia. No biênio 2018/2019, foram coletadas a gravação de 10 interações espontâneas informais em contexto natural. Todas as gravações foram feitas em contexto familiar privado: 6 monólogos e 4 diálogos. Esse plano de trabalho, desenvolvido em 2019 e 2020, tem como objetivo proceder ao tratamento do corpus C-ORAL-Quilombola, realizando a transcrição, revisão e alinhamento texto-som dos dados já coletados. Foi possível realizar o treinamento da equipe transcritora para efetuar a segmentação prosódica das gravações, a partir dos rígidos critérios definidos. Devido à pandemia do novo coronavírus a equipe ficou impossibilitada de retornar a comunidade para novas coletas, cuidando apenas do do material colhido anteriormente. Ao final, realizou-se a transcrição/segmentação de 10 textos, com uma média 1.500 palavras cada, compondo um minicorpus, revisado duas vezes pela equipe, com base em alguns princípios do projeto CHILDES-CLAN, adaptado para o português brasileiro. No entanto, como a fase de transcrição é um empreendimento árduo, tivemos limitação de tempo para iniciar o alinhamento texto-som dos arquivos.

Palavras-chave: CORPUS ORAL, COMUNIDADE QUILOMBOLA, FALA ESPONTANEA, SUL DA BAHIA.