

PELCRA Tools Documentation

This wiki contains descriptions, manuals and documentation for a number of tools and resources developed by the [PELCRA](#) team.

Spokes PL

[Spokes Documentation](#)

The PLLuM Instruction Corpus

[PLLuMIC 1.0](#)

The PLLuM Prompt Book

[PLLuM Prompt Book](#)

SlopeQ for the NKJP

[SlopeQ for NKJP](#)

SlopeQ for the BNC

[SlopeQ for BNC](#)

Spokes for the BNC

[Spokes for BNC](#)

Spokes PL

[Spokes Documentation](#)

Paralela

[Paralela](#)

PELCRA Spoken Offline Corpora

[Spoken Offline Corpora](#)

Treelets

[Treelets](#)

DiaBiz

[DiaBiz](#) is a dialog corpus comprising recordings and annotated transcriptions of phone-based customer-agent interactions in several key business domains.

DiaBiz EN

[DiaBiz EN](#)

DiaBizEval

[DiaBizEval](#)

SpokesBiz

[SpokesBiz](#) is a corpus of conversational Polish developed within the CLARIN-BIZ project and currently comprising over 650 hours of recordings. The transcribed recordings have been diarized and manually annotated for punctuation and casing.

SpokesBiz search engine

[SpokesBiz search engine](#)

SNUV

SNUV (Spelling and NUMbers Voice database) is a spelling and number and recognition speech database containing over 220 hours of recordings of Polish speakers reading numbers and spelling words, recorded in 22050kHz, 16-bit *.wav files. 210 different participants were paid to produce a sample of their speech through an online spoken data collection platform. Written representation of the recordings is provided with the original sound files. The envisaged application of this resource is to enable the creation of automatic speech recognition (ASR) tools that allow users to spell out words and numbers to be recognized. SNUV has been released under a CC-BY license and can be used for both academic and commercial purposes free of charge.

From:

<https://pelcra.pl/> - **PELCRA**

Permanent link:

https://pelcra.pl/doku.php?id=pelcra_tools_documentation

Last update: **2026/07/22 10:05**

