

# HM-Libras Database

## Head movement in Brazilian Sign Language Database

---

### What is the HM-Libras Database?

HM-Libras database of short videos featuring deaf interpreters signing sentences in Brazilian sign language (Libras). The HM-Libras was created with the concept of studying head movement in Libras. For that reason three types of sentences were chosen: affirmative (or neutral), negative and question (wh-question and yes-no question). The facial expression in Libras is common to other sign languages.

HM-Libras is an extension of the QLIBRAS dataset.

The database is composed of 72 FACS annotated and Portuguese translation sequences. There are more than 390 signs with more than 200 different ones among them. All videos were downloaded from the internet under Creative Commons license (CC) and are user-generated content.

User-Generated Content (UGC) videos are uploaded by users and creators. These videos are not always professionally curated and often suffer from perceptual artifacts, illumination, and background variations. For the purpose of this dataset, we've selected original videos with the person facing the camera.

### How to get access to HM-Libras?

Free copies of HM-Libras are provided for research and private study only.  
A copy of the database can be requested through an email sent to:

Emely Pujolli da Silva, MSc  
emelypujolli@gmail.com  
Ph.D. student  
Department of Computer Engineering and  
Industrial Automation  
School of Electrical and Computer  
Engineering  
University of Campinas

Prof. Paula Dornhofer Paro Costa, PhD  
paula@fee.unicamp.br  
Department of Computer Engineering and  
Industrial Automation  
School of Electrical and Computer  
Engineering  
University of Campinas

The e-mail message must contain a description of the envisioned use of the database and a statement of agreement with the following "Terms and Conditions of Use". A download link will be provided after the analysis of the request. Requests can be denied.

# Terms and Conditions of Use

## Copyright

This database is a sampling from a few of UGC videos as uploaded to YouTube distributed under the Creative Commons license. We do not take any ownership of the video content. If you use the content on your own, you agree it can be used under the terms of CC BY 4.0 or any future version of that license. (<https://creativecommons.org/licenses/by/4.0/legalcode>)

All responsibility for possible infringement of copyright or other rights, including the invasion of privacy arising from the use of reproductions is assumed by the user.

Permission is granted for one-time use only. Permission for reproduction and use is limited to the applicant only and is not transferable. The permission for any subsequent change should be requested through the contact above.

This database contains human subjects, to guarantee the proper use of this database all submitted papers (or any publicly available text) that use or talks about the HM-Libras database have to present or print only the set of pictures standard available and strictly designated by the authors.

No economic profit can be made from this database.

## Guidelines on how to refer to HM-Libras in papers and articles.

If this data is used, in whole or in part, for any publishable work, the HM-Libras should be referenced. In general, when mentioning HM-Libras in a publication, please include a reference to one of the following works:

- da Silva, Emely Pujólli, and Paula Dornhofer Paro Costa. "QLIBRAS: A novel database for grammatical facial expressions in Brazilian Sign Language." (2017).
- da Silva, Emely Pujólli, and Paula Dornhofer Paro Costa. "Recognition of non-manual expressions in Brazilian sign language." (2017).

If the image is manipulated in any way, an accompanying caption must indicate that the image has been altered and how it has been altered (significant cropping, manipulation, superimposing, tinting, etc.).

## Files

The Hm-Libras is composed with video and metadata. After your request for access is allowed, you will receive a link for a HM-Libras folder containing three folders and two documents. There are provided two folders with two types of files: Annotation (.xml) and Videos (.mp4). All files are named with coding in the form SiiTFjjVkk where

S - relates to the subject in the video;

TF - associates with the type of phrasing;

V - indicates the video number;  
and here the i,j,k indicates index numbers. For example, file S01TF01V01.xml is the  
annotation for S01TF01V01.mp4 video segment.

## **Research Team:**

Unicamp: Emely Pujolli da Silva, Paula Dornhofer Paro Costa, and José Mario De Martino.  
UFABC: Kate Kumada de Oliveira.