PERMISSION TO DISTRIBUTE TRANSCRIPT DATA THROUGH TALKBANK
Carnegie Mellon University

Carnegie Mellon University is participating in a data-sharing program known as TalkBank that was established through federal and foundation grants to Brian MacWhinney at Carnegie Mellon. For the program to function properly, contributors will be asked to and will give permission for their work to be made available to other researchers. With a full understanding of the aforementioned, I hereby give permission to TalkBank to make and circulate electronic copies of the language transcripts and media that I describe below. These copies may be distributed to scholars and other responsible parties. I warrant that this use of the data is in accord with Human Subjects review procedures at my institution and that participants have given informed consent to have their data available to researchers. I also warrant that there is no copyright restriction over the transcripts and media being circulated. Any further restrictions that I wish to place on the use of these data are listed under (2) below. I do not hold Brian MacWhinney or Carnegie Mellon responsible for the enforcement of these further restrictions and indemnify and render harmless both Brian MacWhinney and Carnegie Mellon University against any actions at law or in equity or in similar courts of any jurisdictions arising from violations of these restrictions.

1. General description of the data set, and IRB Approval # (if available):

EP-Plurals

The research project underlying this corpus is coordinated by Ana Margarida Ramalho and Maria João Freitas, at *Centro de Linguística da Universidade de Lisboa* (CLUL). The corpus was gathered in the context of Ana Margarida Ramalho's PhD project (Ramalho 2017; https://rdpc.uevora.pt/handle/10174/23564). The instrument used for data collection (EP-Plurals) is an appendix of the CLCP-EP (https://clul.ulisboa.pt/recurso/crosslinguistic-child-phonology-project-portugues-europeu), developed in the context of the Cross-Linguistic Child Phonology Project (University of British Columbia), coordinated by Barbara May Bernhardt and Joseph Stemberger.

EP-Plurals (https://www.clul.ulisboa.pt/recurso/ep-plurals) was design to assess the acquisition of European Portuguese (EP) regular and irregular plural forms in the nominal paradigm. It follows the typology in Mateus & Andrade (2000) and includes (the asymmetry in number of items is due to child lexical restrictions):

- 8 items from the regular paradigm (4 words stressed in the penultimate syllable (*Coelho* 'rabbit'); 4 words stressed in the ultimate syllable (*café* 'coffee'))
- 8 items with word-final lateral (*azul* 'blue')
- 8 items with word-final diphthong (*limão* 'lemon')
- 3 items with word-final rhotic (*flor* 'flower')
- 3 items with word-final fricative (*nariz* 'nose')
- 3 items for invariant singular/plural forms (*lapis* 'pencil')

The study is experimental cross-sectional and contains production data from 63 typically developing children from the Lisbon area,), with no history of hearing or language deficits or disorders, all EP monolinguals, aged 3;2 to 6;3 (G1=3;2-3;11; G2=4;0-4;9; G3=5;1-6;3; no audio file and demographic information available for seven children due to lack of parental consent for this information to become public).

The data collection procedures were as follows:

1. informed consents were gathered according to UBC and BREB guidelines;
2. a sequence of 42 digital images from the CLCP-PE test (Ramalho, Freitas, Almeida, Bernhardt & Stemberger, 2014) were presented to the child on an iPad screen for picture naming;
3. a picture was presented and the child was asked to complete the sentence '*This is one…*'; then, a picture with two instances of the same picture was presented and the child was asked to complete the sentence '*These are two…*'.
4. individual audio recordings were made at each child's school.

The Phon software was used (https://www.phon.ca) to perform the phonetic transcriptions. All children were recorded by Ana Margarida Ramalho and transcribed by Mariana Castro. Reliability procedures were adopted. Inter-rater reliability was assessed by having Jéssica Gomes revise 100% of the sample, yielding an agreement rate of 97.1%. In dubious cases, Maria João Freitas carried out a blind transcription; persistent indecision was registered in the database (cf. IPA Actual and Note tiers). Finally, intra-rater reliability was assessed by Jéssica Gomes on 10% of the sample, yielding an agreement rate of 97.7%.

All the phonetic transcriptions and audio files available on Phon are from children whose parents had provided consent.

This research was funded by the Fundação para a Ciência e Tecnologia (FCT, Foundation for Science and Technology) under project numbers UIDB/SFRH/BD 88966/ 2012 (individual PhD grant), UIDB/00214/2020 and UID/214/2025. The project was approved by the UBC Behavioural Research Ethics Board (BREB number H12-00072) and renewed in 2023 by the Research Ethics Committee of the School of Arts and Humanities of the University of Lisbon (FLUL)FLUL (10_CEI2023).

2. Restrictions to be placed on the use of the data:

No restrictions on the use of the data.

Signed  Date 11.11.2026

Printed Name Ana Margarida Ramalho

Institution Centro de Linguística da Universidade de Lisboa/Escola Superior de Saúde do Alcoitão

Please mail this signed form to: Brian MacWhinney, CMU-Psychology, 5000 Forbes Ave. Pittsburgh, PA, 15213, USA or send a scanned signed copy to macw@cmu.edu