Speech Maturity Dataset

William N Havard, Loann Peurey, Kasia Hitczenko, Alejandrina Cristia

Speech Maturity Dataset

William N. Havard, Loann Peurey, Kasia Hitczenco, Alejandrina Cristia

Laboratoire de Sciences Cognitives et Psycholinguistique (ENS–CNRS–EHESS–INRIA–PSL), France

1. Introduction

- **Children's spontaneous vocal productions become increasingly adultlike**
  - **shape** (canonical, non-canonical)
  - **phonetic and phonemic properties** (frequency range, phonotactics, etc.)
- **But... research** has been **limited to a narrow set of languages and communities**
  - **Indo-European languages**
  - **Western**(ised) speaker **communities**
  - **narrow age range: 0 - 24mo**

2. Speech Maturity Dataset

- Superset of **BabbleCor** (Cychosz et al., 2019)
- **15k** (Babblecor) → **258,914 clips** (ours)
- **398 children** (209 male, 186 female)
- Large age range: **2mo - 6yr**
- **14 communities**
  - rich industrialised societies
  - **farmer-forager communities**
- **25+ languages**

3. Zooniverse: Citizen Science

- **Citizen Scientist: Non-scientific volunteers** who annotate and **label scientific data**
- Clip labels based on the **majority vote** of **at least 3 citizen scientists**
- **Majority vote**: **at least 50%** of the citizen scientists **endorsed a particular label**
- **Labels** (speaker type and sex for a subset of the clips only N=110,577)
  - **Vocalisation Type**: canonical, non-canonical, laughing, crying, junk
  - **Speaker Type**: baby (younger than 3 years), child (3-12 years), adolescent (12-18 years)
  - **Sex**: Female/Male (for adolescents and adults)

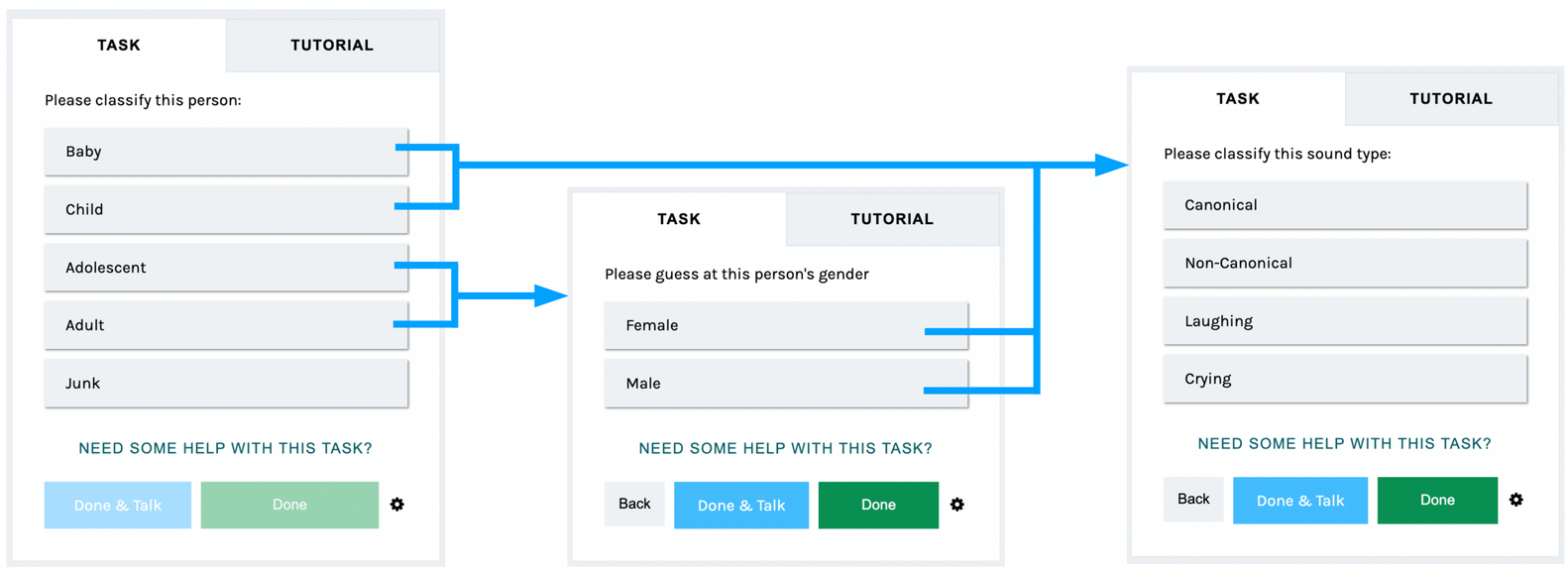

Figure 1: Zooniverse Pipeline

4. Metadata & Use Cases

- **Wealth of metadata**
  - Age
  - Sex
  - Linguistic Environment
  - Normativity
- **Use Cases**
  - Canonical/Linguistic Proportion
  - Train vocalisation-type classifiers

5. Preliminary analysis

- **Canonical (CP)** and **Linguistic proportion (LP)**
  - $\text{CP} = \frac{\text{Canonical}}{\text{Canonical+Non-Canonical}}$
  - $\text{LP} = \frac{\text{Canonical+Non-Canonical}}{\text{Canonical+Non-Canonical+Cry+Laugh}}$
- **Linear Mixed-Effect Model**
  - Predict LP and CP
  - Fixed Effects: age, sex, monolingualism
  - Random Effects: child ID nested in corpus

  ⟶ **Significant positive effect of age**

  ⟶ **No significant effect of age or monolingualism**

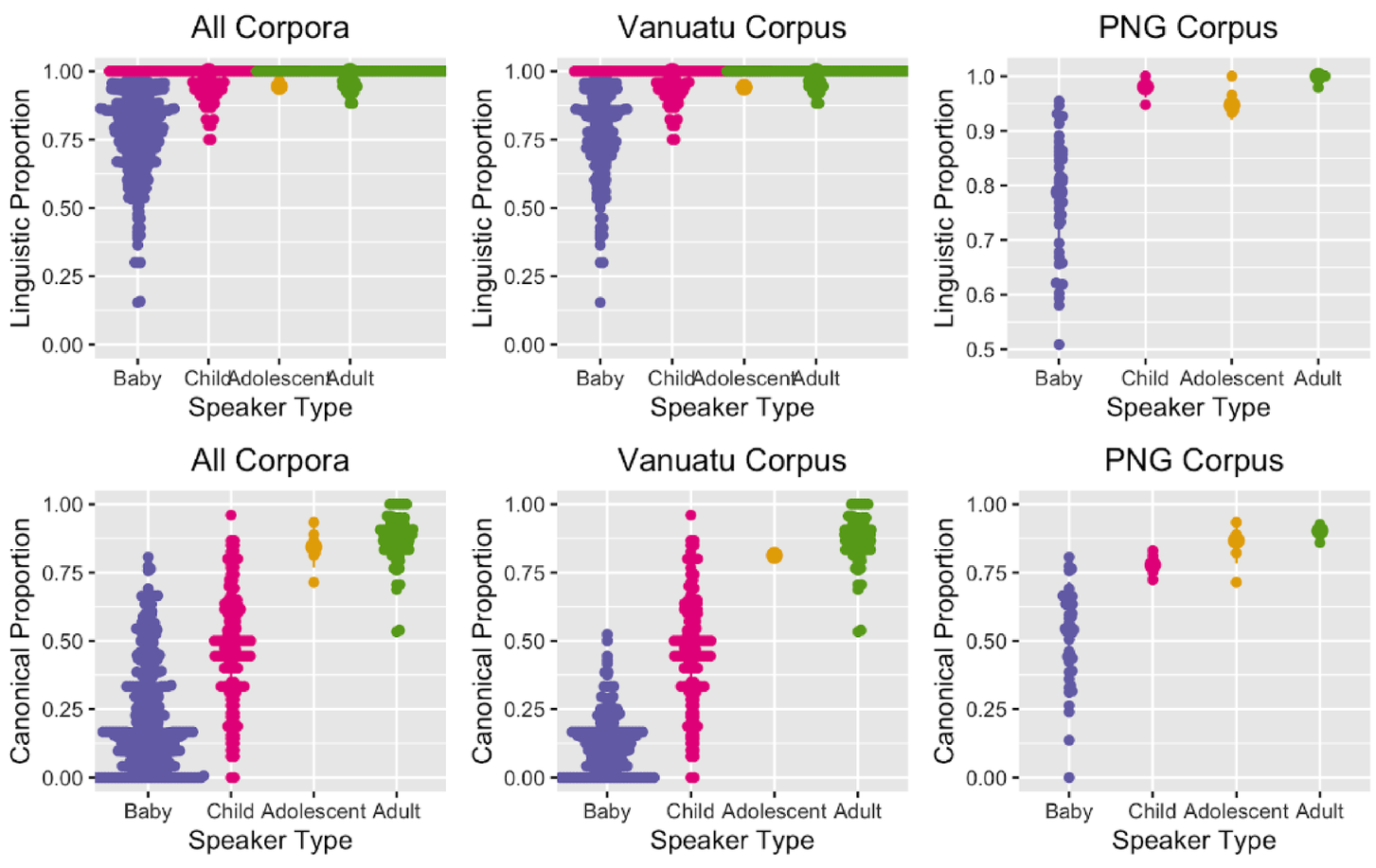

Figure 2: Linguistic proportions (top) and canonical proportions (bottom) by speaker age

This work was supported by the J. S. McDonnell Foundation Understanding Human Cognition Scholar Award; European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (ExELang, Grant agreement No. 101001095), Agence Nationale de la Recherche (ANR-16-DATA-0004 ACLEW). This work used the Extreme Science and Engineering Discovery Environment (XSEDE), which is supported by National Science Foundation grant number OCI-1053575. Specifically, it used the Bridges system, which is supported by NSF award number ACI-1445606, at the Pittsburgh Supercomputing Center (PSC). This publication uses data generated via the Zooniverse.org platform, development of which is funded by generous support, including a Global Impact Award from Google, and by a grant from the Alfred P. Sloan Foundation. The funders had no impact on this study.