*JSLHR* Editorial Board

Paris, November 3, 2020

Dear Editorial Staff,

We are submitting a manuscript entitled *“****Describing vocalizations in young children: A big data approach through citizen science annotation”.*** This paper would be an ideal fit for *Journal of Speech, Language, and Hearing Research* because it provides a unique evaluation of a novel approach to describing young children’s spontaneous vocalizations, as captured via daylong recordings. Although such long-form recordings are increasingly common, challenges remain with respect to how these data are handled, annotated, and analyzed. First, the sheer scale of long-form data means that annotating and describing children’s vocalizations requires a great deal of time. Second, typical assessment of vocal development in young children relies on annotators receiving considerable training, which means data annotation on such a scale becomes not only costly, but also harder to reproduce across laboratories. Our method addresses all of these issues by relying on collaborations with citizen scientists via the Zooniverse platform.

Segments automatically identified by a system called LENA as being produced by the key child were extracted from one daylong recording for each of 20 participants: 10 low-risk control children and 10 children diagnosed with Angelman syndrome, a neurogenetic syndrome characterized by severe language impairments. Speech samples were annotated by trained annotators in the laboratory as well as by citizen scientists on Zooniverse. All annotators assigned one of five labels to each sample: Canonical, Non-Canonical, Crying, Laughing, and Junk. This allowed the derivation of two child-level vocalization metrics: the Linguistic Proportion (proportion of speech-like segments out of all non-Junk segments), and the Canonical Proportion (proportion of Canonical out of all speech-like segments).

Although agreement at the level of individual segments was only moderate, correlations in child-level metrics derived from laboratory and Zooniverse correlated, with all rs above .9. Moreover, reliability did not differ across our two participant groups. We also explored sampling options in our dataset and were able to provide recommendations regarding the quantity of data to be used in such enterprises, and explain the potential implications for the study of speech and language more generally.

We understand it is challenging to find expert reviewers, and thus take the liberty of mentioning the following people, who should be able to provide unbiased evaluations of the present manuscript:

* Jena McDaniel (jena.c.mcdaniel@ku.edu) first-authored a paper on different vocalization metrics extracted from daylong recordings published in *JSLHR* last May
* Laura Dilley (ldilley@msu.edu) coauthored an evaluation as well as a meta-analysis of the LENA system
* Katie Belardi ([KMB326@pitt.edu](mailto:KMB326@pitt.edu)) has done work on canonical babbling ratio and volubility in syndromic populations

We thank you in advance for your time and consideration.

Kind regards,

Alejandrina Cristia, on behalf of all authors