

From: [Karen Mulak](#)
To: [Ping Li](#)
Cc: [Katrina Trewin](#)
Subject: Re: data submission
Date: Friday, 24 March 2017 3:36:46 PM

Thanks, Ping. My responses to your questions are below. Please let me know if anything is unclear, and what the next step would be.

Thank you,
Karen

--

Do you require storage for this data?

Yes, the data would have to be open-access so that anyone could download it.

Do you require a DOI for this data?

Yes

Do you have ethics approval to share this data?

We have submitted an amendment.

What is the title of the dataset? (This should be different to the publication title and be as descriptive as possible)

Dutch- and Australian English-learning 12-month-olds' detection of indexical and linguistic changes to vowel tokens, measured by trial looking time in a serial preference task following familiarization.

Where and when was the data collected (where and when were interviews conducted, GPS coordinates, etc.)

The data were collected at the University of Western Sydney from October 18, 2013 to May 15, 2014, and from the University of Amsterdam from April 4, 2014 to September 23, 2014.

What is the data's temporal coverage?

The data from each participant covered a single test session lasting approximately 3 minutes.

Please provide a description for the dataset. (This should provide readers a complete understanding of what the dataset contains and the conditions under which it was produced. Please make it as detailed as possible).

This data was collected to answer the question of whether infants attend to linguistic and indexical information in vowel perception. Australian-English-learning and Dutch-learning 12-month-olds were played strings of vowel tokens from a central speaker. To measure their attention to the auditory stimulus, A Tobii eyetracker sampling at 120 Hz tracked their attention to a monitor in the vicinity of the speaker. Infants were first familiarized to eight 13-second trials containing repetitions of two tokens of the vowel /ɪ/ spoken by a female speaker of North Holland Dutch (trials Fam1 - Fam8). Infants were then presented with two blocks of three different test trial types, presented in random order within block: NoChange, which presented the same two tokens as in familiarization; VowelChange, in which infants heard the same vowel /ɪ/ alternated with tokens of the vowel /ɛ/ produced by the same speaker; and the IndexicalChange trial, which was either a Speaker Change, meaning that infants heard the same vowel /ɪ/ spoken by the familiarized speaker, alternated with tokens of the vowel /ɪ/ spoken by another female speaker of North Holland Dutch, or an Accent Change, in which tokens of the same vowel /ɪ/ spoken by the familiarized speaker were alternated with tokens

of the vowel /ɪ/ spoken by a female speaker of another accent of Dutch, East Flemish Dutch. Data for each trial are infants' cumulative looking time in ms for each 13-second trial. If infants attended to the linguistic or indexical change in auditory stimulus, we would expect increased looking to that trial relative to the NoChange trial.

Data creators – names in affiliation order please. This may be the same as the manuscript or different.

Karen E. Mulak
Cory D. Bonn
Katerina Chladkova
Richard N. Aslin
Paola Escudero

Are all creators current Western Sydney University staff? If not please provide their affiliation.

Karen E. Mulak

The MARCS Institute for Brain, Behaviour and Development, Western Sydney University

The Australian Research Council Centre of Excellence for the Dynamics of Language

Cory D. Bonn

Department of Brain & Cognitive Sciences, University of Rochester, Rochester, New York, United States

Current Address: Laboratoire Psychologie de la Perception, Université, Paris Descartes, Paris, France

Katerina Chladkova

Amsterdam Center for Language and Communication, University of Amsterdam, Amsterdam, Netherlands

Current Address: Cognitive and Biological Psychology, Institute of Psychology, University of Leipzig, Germany

Richard N. Aslin

Department of Brain & Cognitive Sciences, University of Rochester, Rochester, New York, United States

Paola Escudero

The MARCS Institute for Brain, Behaviour and Development, Western Sydney University

The Australian Research Council Centre of Excellence for the Dynamics of Language

Please advise creators' ORCIDs too if available

Karen E. Mulak: orcid.org/0000-0002-0387-2109

Cory D. Bonn: orcid.org/0000-0001-8795-1682

Katerina Chladkova: orcid.org/0000-0003-2009-1897

Richard N. Aslin: orcid.org/0000-0003-4092-1712

Paola Escudero: orcid.org/0000-0002-8071-7663

Who is the primary contact for the dataset?

Karen Mulak (k.mulak@westernsydney.edu.au)

I will need at least two Keywords to describe the dataset, and if possible a 2/4 or 6 digit FOR code.

170204 Linguistic Processes (incl. Speech Production and Comprehension)
170102 Developmental Psychology and Ageing

Are we able to apply a Creative Commons <<http://creativecommons.org.au/learn/licences/>> CC-BY 4.0 licence which allows reuse with attribution?

No

What size and format is the data – is it one or several files and how big is the dataset (eg 2 Word Documents and 1 Excel spreadsheet 10MG)?

The data file is a 6KB csv file.

What is the Retention period? e.g. perpetual, 5 years, 10 years...

Perpetual

Who owns the data?

The authors own the data.

Was this the result of a research grant? If so I will need details of Funding Bodies, Grant number, Project Title

Australian Research Council Discovery Project DP130102181 "Understanding different speakers vs. different accents: apples and apples or apples and pears?" (CI Escudero)

Australian Research Council Centre of Excellence for the Dynamics of Language (CE140100041; CI Escudero)

Would you like your data to be published in Research Data Australia (RDA) as well? Here's a recently published HIE data set in RDA:

<https://researchdata.ands.org.au/bacterial-microbiome-populations-amplicon-sequencing/820529> <<https://researchdata.ands.org.au/bacterial-microbiome-populations-amplicon-sequencing/820529>>.

Publishing your data in RDA can increase the exposure your data/publication and increase the chance of being cited. (once the related paper is published, I can link the data to the paper).

No

Karen Mulak, PhD | Postdoctoral Research Fellow
ARC Centre of Excellence for the Dynamics of Language
The MARCS Institute for Brain, Behaviour and Development
P: [+61 2 9772 6021](tel:+61297726021) | E: k.mulak@westernsydney.edu.au



From: [Karen Mulak](#)
To: [Ping Li](#)
Subject: Re: data submission
Date: Monday, 27 March 2017 3:52:46 PM
Attachments: [Mulak, Bonn, Chladkova, Aslin & Escudero, 2017.csv](#)

Hi Ping - my additional responses are below, and attached please find the data file.

Thank you,
Karen

On Mon, Mar 27, 2017 at 11:08 AM, Ping Li <P.Li@westernsydney.edu.au> wrote:

- Can you provide a couple of keywords?

infant speech processing, indexical vs. linguistic information, normalization

- What's the type of research activity: Applied research, experimental development, pure basic research, or strategic basic research?

Pure basic research

- You indicated that we couldn't apply CC-BY 4.0 to the data set, allowing reuses with attribution. Can you specify which type of CC licence you would like to apply to this data set: Attribution-Share Alike (CC BY-SA), Attribution-No Derivatives (CC BY-ND), Attribution-NonCommercial (CC BY-NC), Attribution-NonCommercial-Share Alike (CC BY-NC-SA), or Attribution-NonCommercial-No Derivatives (CC BY-NC-ND)? Details can be found here: <http://creativecommons.org.au/learn/licences/>

Sorry, I misread this as reuse without attribution. We can apply reuse with attribution.

Karen Mulak, PhD | Postdoctoral Research Fellow
ARC Centre of Excellence for the Dynamics of Language
The MARCS Institute for Brain, Behaviour and Development
P: +61 2 9772 6021 | E: k.mulak@westernsydney.edu.au

