



Child Directed Print Corpora Understanding the Resource

Virtual Round Table

Co-chairs: Sonali Nag, Jelena Mirkovic & Dina Ocampo
30th June to 2nd July 2021



Web: <http://talktogether.web.ox.ac.uk>

Twitter: @talktogproject

Email: katrina.dulay@education.ox.ac.uk
eve.rogers@education.ox.ac.uk

About the TalkTogether Project

TalkTogether is a multi-country project made possible by an award from the UKRI-GCRF Collective Fund. The wider aim of the project is to develop contextualized early years interventions to support oral language development, and thereby boost children's literacy learning trajectories.

A more immediate aim of the project is to develop the infrastructure needed to support rapid growth in local evidence-building activities that can meaningfully inform intervention development. One such research infrastructure is a child-directed print corpus.

Over 2020, TalkTogether teams developed print corpora in two Asian languages: one in Filipino and one in Kannada.



Background to the Round Table

Print Corpora are useful resources for the study of children's encounters with language in books. The Round Table brings together researchers to discuss how best to examine the print that is likely to shape children's language learning.

Teachers, NGO workers, curriculum developers and government officials from India and the Philippines will also join the meetings.

Print Corpora are used in empirical studies on, for example, the properties of words and sentences and their effects on language and literacy development. Print Corpora may help inform language assessment and language programmes.

The Round Table will examine these uses of Print Corpora, and also how language in books is unique and different from spoken language.

Objectives

To understand the promise of child-directed print corpora for experimental studies and child language assessment.

To understand how child-directed print corpora can support theorizing on morphological and morpho-syntactic development, usage and input-based approaches to child language acquisition, and language processing.

To understand the usefulness of child-directed print corpora for practitioners and their trainers, and curriculum developers.

These themes will be reflected in three interlinked sessions on Child Language Assessment, Experimental Research and Children's Materials.



UTC	Madison	Toronto	Oxford	Bangalore	Manila	Malaysia	Melbourne
6.30-9.00	1.30-4.00	2.30-5.00	7.30-10.00	12.00 -14.30	14.30-17.00	14.30-17.00	16.30-19.00

ROUND TABLE 1

Child Language Print Corpora and Child Language Assessment

CHAIR: Sonali Nag

DATE: 30th June

INTRODUCTION

Oral Presentations

- **Cross-linguistic lexical properties of child-directed print corpus: Implications for Filipino and Kannada assessments**
John, S., Agarwal, A., Ocampo, D., Tiwari, S., Lucasan, K. L. & Nag, S.
- **Implications for morphological awareness for Filipino Language and literacy assessment.**
Bacolod, M.A., Javier, J., Padilla, P., Ocampo, D., Sicat, C., Gonzales, J., Lucasan, K., Perez, A., Rosario, F. & Nag, S.
- **Structural properties of child-directed print corpus in Kannada: Implications for child language assessments**
Tiwari, S., Nakshathri, M., Agarwal, A., John, S. & Nag, S.

BREAK

Panel Discussion

- **Research and applications of Malay Language Word Corpus**
Lee, L.W. & Low, H.M.
- **Morphology and the development of orthography-meaning links: A large-scale corpus study of children's books**
Dawson, N.

Discussant

Parrila, R.



UTC	Madison	Toronto	Oxford	Bangalore	Manila	Malaysia	Melbourne
11.30-14.00	6.30-9.00	7.30-10.00	12.30-15.00	17.00-19.30	19.30-22.00	19.30-22.00	21.30-24.00

ROUND TABLE 2

Child Language Print Corpora and Experimental Research

CHAIR: Jelena Mirkovic

DATE: 1st July

INTRODUCTION

Oral Presentations

- **The development of complex syntax in children's reading and writing**
Hsiao, Y., Dawson, N., Banerji, N. & Nation, K.
- **Children's acquisition of a symmetrical voice language: Evidence from comprehension and production experiments in Tagalog**
Garcia, R.
- **Which word properties are related to age-of-acquisition ratings in Kannada and Filipino?**
Dulay, K., Kotian, D., Fua, M.M.R.C., Nakshathri, M., Reoperez, M., Nagendra, S., Haydee, E., Alipustain, G., Pandith, P., Ocampo, D., Padilla, P. & Nag, S.
- **Reading the mind: Fiction reading supports emotion recognition**
Schwering, S.

BREAK

Panel Discussion

- **Writing vs. Speaking: How production goals influence language, with implications for corpus analyses**
MacDonald, M.C.
- **Learning, processing and reading development: lessons and impressions from corpus explorations**
Nation, K.

Discussant

Lieven, E.



UTC	Madison	Toronto	Oxford	Bangalore	Manila	Malaysia	Melbourne
11.30-14.00	6.30-9.00	7.30-10.00	12.30-15.00	17.00 -19.30	19.30-22.00	19.30-22.00	21.30-24.00

ROUND TABLE 3

Child Language Print Corpora and Children's Materials

CHAIR: Dina Ocampo

DATE: 2nd July

INTRODUCTION

Oral Presentations

- **Folk-Inspired Kannada Intervention: The Promise Audio Corpus**
Kuchibhotla, S. & Arulmani, G.
- **Of heights and limits: A children's book levelling experience**
Padilla, P., John, S., Nagendra, S., Dulay, K., Concepcion-Alcazar, M.Y., Diaz, L., Ongtengco, M.H., Reoperez, M.G. & Nag, S.
- **Readers and text complexity: an emerging framework for assessment**
Diaz, L., Lucasan, K., Fua, M.M.R.C., & Ocampo, D.

BREAK

Panel Discussion

- **Child-directed text corpora: A great start**
Montag, J.
- **The Spontaneous Spoken Language of Filipino Preschool Children**
Marzan, J.

Discussant

Peterson, S.S.



About Session Chairs



Mirkovic, Jelena is Senior Lecturer at the School of Education, Language, and Psychology at York St. John University. She is interested in the cognitive and neural mechanisms underpinning language learning and use. She has examined morphologically complex languages using computational and behavioral methods, including how they are shaped by domain-general learning and memory processes. She leads the Cognitive Psychology Research Group at York St. John University.



Nag, Sonali is Associate Professor at the Department of Education and a fellow of Brasenose College, University of Oxford. She is the Principal Investigator of the TalkTogether Project. Sonali investigates child learning within diverse settings. Her research is comparative with a focus on languages, writing systems, cultural settings and levels of socio-economic status. Her work seeks to develop a nuanced, contextually grounded understanding of child development.



Ocampo, Dina is Professor at the College of Education, University of Philippines where she served as the Dean of the College of Education from 2010-2013. Her research focuses on literacy development and difficulties among Filipino-English bilingual children. She served as the Undersecretary for Curriculum and Instruction at the Department of Education for the Republic of Philippines from 2013-2017. She is a co-investigator of the TalkTogether Project.



About Discussants



Lieven, Elena is Emeritus Professor at the Division of Human Communication, Development and Hearing, University of Manchester and a Fellow of the British Academy. Her research examines the link between linguistic input and language development, the environmental differences in input characteristics, and the development of grammar. She was instrumental in setting up the Max Planck Child Study Centre. Recently, she led the ESRC International Centre for Language and Communication Development during Phase 1.



Parrila, Rauno is Professor at the Department of Educational Studies, Macquarie University. He is an active member of the Centre for Reading. His research is in Educational Psychology, Developmental Psychology and Cognitive Psychology. His current research examines the impact of reading difficulties and their compensation in different languages and education levels. His body of work is cross-linguistic covering languages of Europe and Asia.



Peterson, Shelley Stagg is Professor at the Department of Curriculum, Teaching and Learning, University of Toronto-Ontario. Her research interests are wide ranging: collaborative action research, rural and indigenous education, play-based learning, child assessment and children's literature. She has also served as the President of the Toronto Reading Council and is a former Board Member of the International Literacy Association.



About Speakers



Bacolod, Mary Ann is Associate Professor at the Department of Linguistics, University of the Philippines. She has held several key positions within and outside the University: Chairperson, The Technical Committee for Foreign Languages; Chairperson, UP Department of Linguistics; Consultant, Bureau of Immigration Training. Within TalkTogether, she is involved in Part-of-Speech tagging and analysing the Affix-Mapping of the Child-Directed Corpus for the Filipino language.



Dawson, Nicola is a Post-Doctoral Research Associate at the Department of Experimental Psychology, University of Oxford. Her research seeks to understand the influence of spoken and written linguistic experiences in shaping language development. In her current work, she is exploring how morphological processing in reading develops based on person and item level factors.



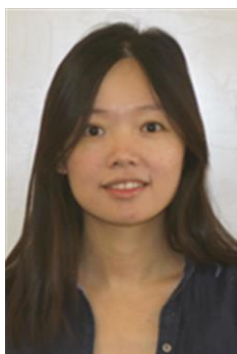
Diaz, Leonor is Assistant Professor at the College of Education, University of Philippines. Her research interests include play and early childhood, text assessment, and early literacy. As part of the TalkTogether, she was involved in the levelling of children's books according to age band.



Dulay, Katrina is a Research Officer in the TalkTogether Project at the Department of Education, University of Oxford. She is interested in analysing family and individual variables influencing language, literacy and numeracy skills among Filipino children. She has developed and monitored the impact of family-based interventions to enhance these skills.



Garcia, Rowena is a post-doctoral researcher at the Max Planck Institute for Psycholinguistics, Netherlands and is interested in researching the acquisition of Tagalog as a first language. She is currently investigating thematic role assignment and word order preference in Tagalog. She is a speech and language pathologist by training.



Hsiao, Yaling is a British Academy post-doctoral researcher at the Department of Experimental Psychology and a fellow of Wolfson College, University of Oxford. Her research examines whether certain experiences result in easier and faster processing of grammar and words.



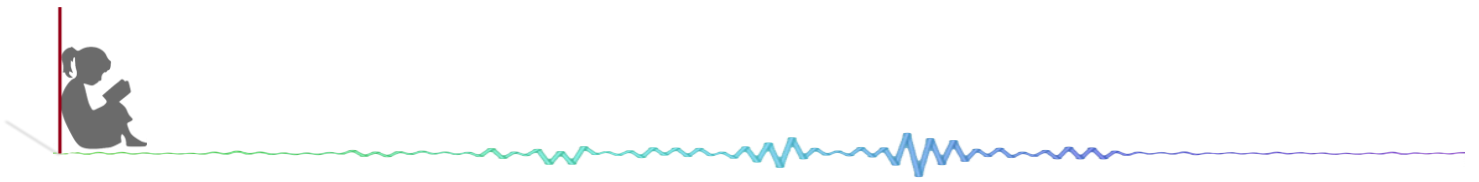
John, Sunila is Associate Professor at Department of Speech & Hearing, Manipal College of Health Professions, Manipal Academy of Higher Education. She is part of TalkTogether in the capacity of Research Affiliate, and is involved in the psycholinguistic analysis of the print corpora developed over 2020.



Kuchibhotla, Shriya is a Materials Development Officer at The Promise Foundation. Within the TalkTogether Project, she is leading the development of folk-inspired classroom interventions for pre-school children. She received her Master's in Philosophy in Education (Critical Approaches to Children's Literature), from the University of Cambridge and has also worked as a primary school teacher in Bangkok, Thailand.



MacDonald, Maryellen is a Donald P. Hayes Professor of Psychology and Associate Chair for Undergraduate Studies, University of Wisconsin-Madison. Her research examines the role of experience and working memory in language comprehension. She pioneered the Production-Distribution-Comprehension Theory which outlines the role of language production in language comprehension. She is also a Co-Director of the Language and Cognitive Neuroscience Lab.



Marzan, Jocelyn is Associate Professor at the Department of Speech Pathology, University of Philippines. She has developed teaching and learning materials on a wide range of topics: normal communication development, communication and swallowing disorders as well as role of communication competence in development. As member of the Board of Trustees of the Philippines Association of Speech Pathologists, she was pivotal in lobbying for the passing of the Speech Language Pathology Act (2016).



Min, Low Hui is Lecturer at the School of Educational Studies, University Sains Malaysia. Her research interests are wide ranging and encompass: augmented and alternative communication, childhood multilingualism and speech and language pathology.



Montag, Jessica is Assistant Professor at the Department of Psychology, College of Liberal Arts and Sciences, University of Illinois Urbana-Champaign. Her research analyses the environmental factors that influence the developmental trajectory of language abilities. She is the Principal Investigator at the Learning and Language Lab and an affiliate at the Beckman Institute for Advanced Science and Technology.



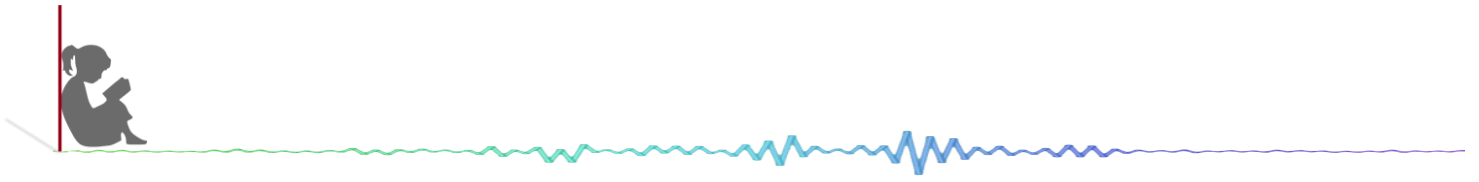
Nation, Kate is Professor at the Department of Experimental Psychology and a fellow of St. John's College, University of Oxford. Her research aims at uncovering the processes underlying reading development, specifically those involved in decoding and comprehension. She has received several awards for her work including the Spearman Medal, the Experimental Psychology Society Prize and the Eminent Researcher Award. She is the Director of ReadOxford and the Language and Cognitive Development Research Group.



Padilla, Portia is Associate Professor at the College of Education, University of Philippines. Her research interests include bi/multilingual reading, children's literature, and helping struggling readers. She is a co-investigator of the TalkTogether Project.



Schwering, Steve is a Graduate student working in the Language and Cognitive Neuroscience Lab, University of Wisconsin-Madison. His research explores the relationship between language and memory. In particular, he is interested in the syntactic choices individuals make to convey meaning and the memory demands underlying these choices.



Tiwari, Shivani is Associate Professor at the Department of Speech and Hearing, Manipal College of Health Professions, Manipal Academy of Higher Education. Her research interests span across the spectrum of typical and atypical development. She is a Research Affiliate of TalkTogether and is involved in the psycholinguistic analysis of the Kannada print corpus.



Lee, Lay Wah is currently Professor at the School of Educational Studies, USM. Before this, she was the Chairperson of the Special Education Program. She is the curriculum developer of both the B.Ed (Special Education) and the forthcoming M.Ed (Special Education) program. Her research focuses on topics including dyslexia, learning disabilities, special educational technology and visual impairment.



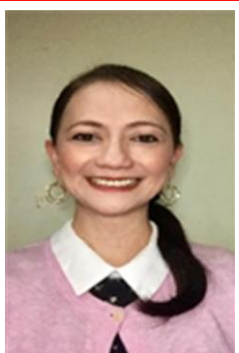
About Co-authors



Agrawal, Aakash is working with Prof. Stanislas Dehaene at Neurospin, Paris, to develop deep neural network models that predict the neural underpinnings of reading. His research on language and vision has won him several awards including the Institute medal for the best PhD thesis in BSSE. Within TalkTogether, he is developing algorithms to analyse large text corpus and is involved in automating child assessment using web-based apps.



Arulmani, Gideon is the Director of The Promise Foundation, India. He is a clinical psychologist by training who is interested in culture and counselling. He has held key leadership positions in several national and international organizations: Founder-Trustee, Indian Association for Career and Livelihood Planning; Board Member, IAEVG; International Fellow, NICEC. His Cultural Preparation Process Model has informed intervention design in more than 30 countries. He is a co-investigator at TalkTogether, representing India through The Promise Foundation.



Concepcion-Alcazar, Marie Yvette is Assistant Professor at the College of Education, University of Philippines. Before joining the college, she taught Kindergarten and early elementary classes at the UP Integrated School. Her research interests include language use in counseling, counseling young children, early literacy, and play. As part of the TalkTogether Project, she was involved in the leveling of children's books according to age band



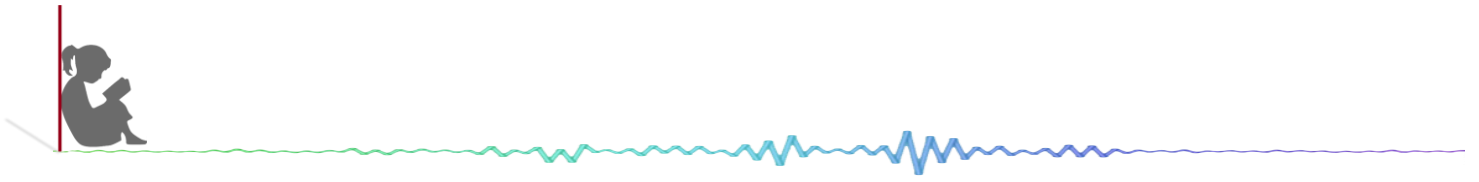
Fua, Margaret Mary Rosary Carmel is Lecturer at the College of Education, University of Philippines. She is an early childhood development and literacy worker. Her research interests include early literacy, with a specific focus on writing and literacy assessment. As part of the TalkTogether, she led the study on the age of acquisition and visualisation of words from children's books.



Gonzales, Junette Fatima is a Senior Research Associate at the University of the Philippines as part of the Center for Integrative and Development Studies Education Research Program. Her research interests include flexible learning options and ICT in education. Within TalkTogether, she is involved in part-of-speech tagging and affix-mapping of the child-directed corpus for the Filipino language.



Javier, Jem R. is Assistant Professor at the Department of Linguistics, College of Social Sciences and Philosophy, University of the Philippines. He has an MA in Linguistics and conducts research primarily on Philippine linguistics and culture studies, as well as on the interface between language and cognition. Currently, he is involved in the UP Monolingwal na Diksiyonaryong Filipino project, based on the digital text corpus of the contemporary use of Filipino. Within TalkTogether, he is involved in the affix-mapping of the child-directed corpus for the Filipino language.



Kotian, Deekshitha is a Research Associate at The Promise Foundation, India. Her research examines children's oral language development. In the TalkTogether project, she worked on the age of acquisition study and morphological analysis of sentence corpus.



Lucasan, Kathrina is a Senior Research Associate at the University of the Philippines as part of the Center for Integrative and Development Studies-Education Research Program. Her research interests include literacy assessment and multilingual education. Within TalkTogether, she is involved in the part-of-speech tagging and affix-mapping of the child-directed corpus for the Filipino language.



Nagendra, Sanjana is a Technical Project Coordinator at The Promise Foundation, India. As part of the TalkTogether Project, she led the creation of the Print Corpus in Kannada which involved selection, digitisation and levelling of the books according to age band.



Nakshathri, Megha is a Research Associate at The Promise Foundation, India. Her research examines children's oral language development. In the TalkTogether project, she worked on the morpho-syntactic analysis of the sentence corpus and was part of the Age of Acquisition survey which was administered to teachers, language experts, and parents to understand the profile of children's language development in Kannada.



Ongtengco, Maria Hazelle is Assistant Professor at the College of Education, University of Philippines. Her research interests include instructional and assessment materials development for literacy teaching and learning, the use of software in developing the literacy skills of learners, the role of children's literature in literacy instruction, acknowledging voices in the literacy classroom, and equity in literacy education. As part of the TalkTogether Project, she was involved in the leveling of children's books according to age band.



Perez, April is Assistant professor at the Department of Filipino and Philippine Literature, College of Arts and Letters, University of Philippines. Her research interests include Filipino linguistics, structural linguistics, sociolinguistics, corpus linguistics, and teaching Filipino language to non-native speakers. Within TalkTogether, she is involved in part-of-speech tagging and affix-mapping of the child-directed corpus for the Filipino language.



Reoperez, Marie Grace is Assistant Professor at the College of Education, University of Philippines. Her research interests include adolescent and adult literacy, content area literacy instruction and assessment, and digital literacy. As part of the TalkTogether Project, she was involved in children's book leveling and the study on the age of acquisition and visualisation of words from children's books.



Rosario, Francisco Jr. is Assistant professor at the Department of Linguistics, University of the Philippines. He conducts research primarily on the Pangasinan language and other Philippine languages, as well as on the interface of language, culture, and society. He served as a semanticist for the UP Monolingual Dictionary Project, and a language consultant for data annotation, transcription, and localization for different Tagalog Text-to-Speech (TTS) projects. Within TalkTogether, he is involved in part-of-speech tagging of the child-directed corpus for the Filipino Language.



Sicat-De Laza, Crizel is Assistant Professor at the College of Education, University of Philippines. Her research interests include Filipino language teaching and learning. Within TalkTogether, she is involved in part-of-speech tagging and affix-mapping of the child-directed corpus for the Filipino language.