INVITED COMMENTARY

Collecting language acquisition data from understudied urban communities: A reply to Cristia et al.

Rowena GARCIA[1,2,*] , Hannah Maria D. ALBERT[3] , Ivan Paul BONDOC[4] and Jocelyn Christina B. MARZAN[3]

[1]Department of Linguistics, University of Potsdam, Germany
[2]Language Development Department, Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands
[3]Department of Speech Pathology, University of the Philippines Manila
[4]Department of Linguistics, University of Toronto, Canada
*Corresponding author: Department of Linguistics, University of Potsdam, Haus 14, Karl-Liebknecht-Str. 24-25, 14476 Potsdam, Germany. Email: garcia1@uni-potsdam.de

In the target article, Cristia, Foushee, Aravena-Bravo, Cychosz, Scaff, and Casillas (2023) convincingly show the need to broaden the current language acquisition research base, not only in linguistic diversity, but also in terms of regions and cultural groups studied. In conducting acquisition research in understudied populations, such as in rural settings, the authors highlight the importance of using a multi-method approach. They present the challenges in adapting these methods to new settings and offer possible ways to promote this type of research. In this commentary, we extend the discussion to understudied urban communities, as we encounter several of the concerns raised in Cristia et al. when collecting observational and experimental language acquisition data from Metro Manila, Philippines. We first describe the community we study, the challenges and modifications needed for conducting research in this setting, and end with a discussion of possible strategies to promote research in communities with understudied populations.

The field context

Most of our work focuses on children's acquisition of Tagalog, an understudied Western Austronesian language with more than 25 million speakers (Eberhard, Simons & Fennig, 2022). In terms of language acquisition research, Tagalog is of interest because it has a unique voice-marking system where the affix on the verb assigns the thematic role of the subject-like argument (Himmelmann, 2005). Aside from using Tagalog's properties to test the predictions of different acquisition theories, as trained speech pathologists, we also aim to establish language development norms for our community to better serve our clients. Despite the high number of speakers, research on Tagalog language acquisition has been scarce.

CrossMark

Aside from Tagalog, the majority of the population have significant exposure to English and/or another Philippine language through media, educational or work settings – hence, there is a need for some measure of language dominance of the families prior to data collection (Marzan, 2009). Families who mostly use Tagalog usually come from lower socio-economic backgrounds, as richer or more highly educated guardians currently prefer to use English (Amora, Garcia & Gagarina, 2020). The families we have visited for our naturalistic studies (Garcia & Kidd, 2022; Marzan, 2013) live in small houses (e.g., living room is also the bedroom) located in tiny alleys, with minimal (safe) space for children to play. Similar to rural communities, many participants live with extended family members – hence, childcare and language input are most likely distributed across several family members. In contrast to the communities described in Cristia et al., our participants have access to electricity, telecommunications, and industrial products, although their access to health services and education is somewhat limited.

Opportunities and challenges from designing the study to collecting the data

When creating experimental design and stimuli, we agree with the authors that it is important to collaborate extensively with native speakers and locals. It is to our advantage that we are ourselves native speakers of Tagalog and usually members of the communities we study. Moreover, there is considerable linguistic work on Tagalog already available for use. However, extensive annotated Tagalog corpora (for calculating frequency statistics) and communicative development inventories (for determining age of acquisition) to inform stimuli creation do not yet exist.

Another benefit of being locals is that with our shared language, customs, and values, fellow Filipinos more readily decide to participate in our research. In the few instances when we are not part of a particular community, recruitment is expedited by establishing a relationship with someone from the target community who then invites neighbours to participate; as also suggested in the target article. Coming from the same culture, we know that some guardians prefer to listen to a simple explanation of the study instead of reading the informed consent documents. Some also prefer answering questionnaires verbally rather than filling out forms. Some parents also volunteer their children without asking for the children's assent, which is consistent with the cultural position of children in the Filipino family (Alampay, 2014). In these cases, we seek the children's assent ourselves, taking the time needed to develop rapport and discuss the study with them, or provide a video explaining what the study entails.

Working with low-income communities, we try to ensure that any compensation we provide is not seen as coercive. Alternative non-financial compensation, such as providing books to schools, offering primer information on speech and language development to parents, story-book reading for the community, and screening children for speech impairments are received well by participating families and schools. This is also a form of giving back to the community rather than merely taking data from them.

Our urban setting differs from the rural setting in that population size is not small. We typically find participants for data collection in a single community. For example, in Garcia, Garrido Rodriguez, and Kidd's (2021) experiment, 154 child participants were all from one public school. This also means that data collection can proceed quickly. When recruiting infants or toddlers, however, one cannot rely on public institutions like daycare centers. Nor is there an acquisition lab in the Philippines with a working database of potential participants.

Regarding participant selection, similar to rural communities, identifying typically-developing children is challenging for us because we lack established norms and stand-ardized assessment tools. We usually rely on teachers' judgement of children's similarity to peers. Our clinical background also allows us to determine who among the children might show atypical development (Marzan, 2013; Marzan, Cabrera, Cunanan, Deleña, Javier & Narcida, 2017). However, involving trained clinicians might not be possible in many communities, so the best way would be to get an indication from people who know the children and can compare them with their peers.

Conducting research in an urban community setting also entails challenges during the testing sessions per se. Since we do not have established labs for linguistic and related research (as academic institutions have just started transitioning from teaching to full-fledged research universities), we collect data in someone else's office, at the back of classrooms, and also in informal settlements (e.g., Bondoc, O'Grady, Deen & Tanaka, 2018a; Bondoc, O'Grady, Deen, Tanaka, Chua, De Leon & Siscar, 2018b; Bondoc, Deen, Or & Hemedes, 2019; Tanaka, O'Grady, Deen & Bondoc, 2019; Tanaka, Bondoc & Deen, 2022) – hence, there are several distractors that we cannot control. Typically, noise levels are high and space is limited. To compensate, we use headphones, multiple recorders, and make written notes. Such an experimental task set-up is usually new to the participants – hence, ample time for practice is provided. Weather is another factor: Philippine heat and humidity are conducive to the breakdown of gadgets, so back-up options are needed. For example, in Garcia, Roeser, and Höhle's (2020) eye-tracking study, a webcam also recorded the session in case manual annotation would be needed, and an extra monitor was used to view where the participants were looking in real time. On occasion, available testing areas are dark, electric outlets are limited, or roofing and furniture are improvised. In these cases, we opt for experiments which do not require electricity (Garcia & Kidd, 2020). We prepare artificial light sources and additional batteries. We also do not indicate to participants that the conditions are not ideal, to not make them feel shy about the set-up. In a recent study, Albert (2022) collected data online, which had its own challenges. Internet speed and faulty gadgets were an issue, as well as participants' lack of familiarity with the setup. When parents were informed about the level of supervision needed from them, data collection was facilitated. In general, we tend to recruit additional participants to compensate for the non-lab conditions of our data collection.

Increasing research on understudied populations

We fully endorse the target article's proposal to make the peer-review system more flexible with regard to sample sizes and methodological variations, and to not require "control" groups from over-represented communities. We agree that conducting research in understudied populations should be more incentivized, as discussed in Kidd and Garcia (2022), in view of the challenges involved.

We also support the suggestions of Cristia et al. regarding increasing researcher diversity. We hope to see more training and career development opportunities for members of under-represented groups. For our group in particular, the opportunity to study, train and develop networks outside of the Philippines has enabled us to bring what we learned back to our community. We also encourage the establishment of more collaborations between international universities with substantial resources and local universities, which are typically not so well-resourced. From our experience working in a university in the Global South, we can say that academics in under-represented

communities not only have the issue of limited funding for research, but might have administrative and/or teaching loads that preclude extensive collaboration or initiation of their own research projects. Changes in the local universities' policies and structures are needed. Extending the authors' proposal to collaborate with experts in ethnographic methods, we would also like to suggest collaborating with other professionals in the target area, such as psychologists and speech pathologists. Moreover, we recommend collaborating with the local government or non-governmental organizations on research projects which have direct impact to the community.

Lastly, we hope that the main message of the target article – the need to diversify our language acquisition research base – will reach many researchers, reviewers and editors, and employment and funding bodies, as it will take a village for us to change the current state of our field.

References

Alampay, L. P. (2014). Parenting in the Philippines. In H. Selin and P. Schvaneveldt (Eds.), *Parenting across cultures: Childrearing, motherhood and fatherhood in non-Western cultures* (pp. 105–121). The Netherlands: Springer.

Albert, H. M. D. (2022). *Tagalog sentence repetition test for Tagalog speakers: Content validation and pilot testing* [Unpublished master's thesis]. University of the Philippines Manila.

Amora, K. K., **Garcia, R.**, & **Gagarina, N.** (2020). Tagalog adaptation of the Multilingual Assessment Instrument for Narratives: History, process and preliminary results. *ZAS Papers in Linguistics*, **64**, 221–233.

Bondoc, I. P., **O'Grady, W.**, **Deen, K.**, & **Tanaka, N.** (2018a). *Effects of pronoun case on sentence comprehension among Tagalog children* [Poster presentation]. The 25th Meeting of the Austronesian Formal Linguistics Association, Taipei, Taiwan.

Bondoc, I. P., **O'Grady, W.**, **Deen, K.**, **Tanaka, N.**, **Chua, E. C.**, **De Leon, A. C.**, & **Siscar, J. A.** (2018b). More relativization asymmetries: Children find locative and benefactive clauses difficult. In A. B. Bertolini & M. J. Kaplan, *The Proceedings of the 42nd Annual Boston University Conference on Language Development* (pp. 72–85). Somerville, MA: Cascadilla Press.

Bondoc, I. P., **Deen, K.**, **Or, E. M.**, & **Hemedes, M. C.** (2019). Reflexives in adult and child Tagalog. In M. M. Brown & B. Dailey (Eds.), *The Proceedings of the 43nd Annual Boston University Conference on Language Development* (pp. 82–93). Somerville, MA: Cascadilla Press.

Cristia, A., **Foushee, R.**, **Aravena-Bravo, P.**, **Cychosz, M.**, **Scaff, C.**, & **Casillas, M.** (2023). Combining observational and experimental approaches to the development of language and communication in rural samples: Opportunities and challenges. *Journal of Child Language*, **50**, 495–517, https://doi.org/10.1017/S0305000922000617

Eberhard, D. M., **Simons, G. F.**, & **Fennig, C. D.** (Eds.) (2022). *Ethnologue: Languages of the World* (25th edition). Dallas, Texas: SIL International. Online version: http://www.ethnologue.com

Garcia, R., **Garrido Rodriguez, G.**, & **Kidd, E.** (2021). Developmental effects in the online use of morphosyntactic cues in sentence processing: Evidence from Tagalog. *Cognition*, **216**: 104859.

Garcia, R., & **Kidd, E.** (2020). The acquisition of the Tagalog symmetrical voice system: Evidence from structural priming. *Language Learning and Development*, **16**(4), 399–425.

Garcia, R., & **Kidd, E.** (2022). Acquiring verb-argument structure in Tagalog: A multivariate corpus analysis of caregiver and child speech. *Linguistics*, **60**(6), 1855–1906.

Garcia, R., **Roeser, J.**, & **Höhle, B.** (2020). Children's online use of word order and morphosyntactic markers in Tagalog thematic role assignment: An eye-tracking study. *Journal of Child Language*, **47**(3), 533–555.

Himmelmann, N. P. (2005). Tagalog. In K. A. Adelaar, & N. P. Himmelmann (Eds.), *The Austronesian languages of Asia and Madagascar* (pp. 350–376). London, England: Routledge.

Kidd, E., & **Garcia, R.** (2022). Where to from here? Increasing language coverage while building a more diverse discipline. *First Language.* Advance online publication.

Marzan, J. B. (2009). *Bilingual language development in the Philippine context* [Unpublished manuscript]. University of the Philippines Diliman.

Marzan, J. B. (2013). *Spoken language patterns of selected Filipino toddlers and pre-school children* [Unpublished doctoral dissertation]. University of the Philippines Diliman.

Marzan, J. B., **Cabrera, L. S.**, **Cunanan, D. R.**, **Deleña, M. O.**, **Javier, Y. L.**, & **Narcida, K. C.** (2017). *Content-form interactions of selected Filipino children aged 14-to-60 months* [Unpublished manuscript]. University of the Philippines Manila.

Tanaka, N., **O'Grady, W.**, **Deen, K.**, & **Bondoc, I. P.** (2019). An asymmetry in the acquisition of relative clauses: Evidence from Tagalog. *First Language*, **39**(6), 618–632.

Tanaka, N., **Bondoc, I. P.**, & **Deen, K.** (2022). Examining main clause similarity and frequency effects in the production of Tagalog relative clauses. *Journal of the Southeast Asian Linguistics Society*, **15**(2), 70–86.