

Author response to reviews of

Manuscript LING.2023.0249

Opening open science to all: Demystifying reproducibility and transparency practices in linguistic research

Joseph V. Casillas, Gabriela Constantin-Dureci, Iván Rascón Andreu, Jiawei Shao, Stephanie A. Rodríguez, Adrija Gadamsetty, Alexandria Minetti, Krishita Laungani, John Thatcher, Rhode-Taina Gardere, Katherine Taveras, Isabelle Chang, Nicole Rodríguez, Meritxell Feliu Ribas & Kyle Parrish
submitted to *Linguistics: An Interdisciplinary Journal of the Language Sciences*

[RC] Reviewer comment

Manuscript text

Dear Dr. Volker Gast,

Thank you for taking the time to consider our manuscript *Opening open science to all: Demystifying reproducibility and transparency practices in linguistic research* () for publication in *Linguistics: An Interdisciplinary Journal of the Language Sciences*. Our understanding is that we needed to consider the reviewers' comments and revise accordingly before the manuscript could be reevaluated for publication. We noted X specific areas that were consistently highlighted by the reviewers as needing improvement: x, y, and z. We have considered thoroughly the detailed feedback provided by all three reviewers and resubmit what we believe to be a much improved version of the manuscript. In this letter we address the reviewers' concerns point-by-point. Where feasible, we quote all revised text in this document, otherwise, we refer the reader to the relevant sections of the revised manuscript. We thank you and the three anonymous reviewers for all comments and suggestions and again enthusiastically submit the revised manuscript for consideration in *Linguistics: An Interdisciplinary Journal of the Language Sciences*.

Sincerely,
Joseph Casillas
(corresponding author)

1. Reviewer #1

[RC 1.1.] Let me begin by first clarifying my own positionality with regard to the subject at hand: I am an enthusiastic supporter of open science practices in linguistics, and have been for many years. So in this sense, I am solidly in favor of more exhortations in print to our colleagues to improve their practices, and recognize that many linguists are not trained in these practices and left to more or less fend for themselves. This is a sad state of affairs for our field and others, and needs to be corrected through outreach and education.

This paper, however, seems to me to be quite weak on two fronts. The first is what seems to be a concerning lack of awareness of recent work in linguistics to promote open science, especially with regard to research data. Of particular note is the absence of two fairly recent references in linguistics. The first is a highly-cited, multi-authored position paper calling for reproducible linguistics research that was published in this very journal in 2018:

Berez-Kroeker, Andrea L., Lauren Gawne, Susan Kung, Barbara F. Kelly, Tyler Heston, Gary Holton, Peter Pulsifer, David Beaver, Shobhana Chelliah, Stanley Dubinsky, Richard P. Meier, Nicholas Thieberger, Keren Rice & Anthony Woodbury. 2018. Reproducible research in linguistics: A position statement on data citation and attribution in our field. *Linguistics* 56(1): 1–18. <https://doi.org/10.1515/ling-2017-0032>.

The second is an open access handbook from the very prominent MIT Press Open, with 56 chapters on open data management from more than 100 international authors across many subfields (which I refer to later in this review as OHLDM):

Berez-Kroeker, Andrea L., Bradley McDonnell, Eve Koller & Lauren B. Collister (eds.). 2022. The open handbook of linguistic data management. Cambridge, MA: MIT Press Open. <https://doi.org/10.7551/mitpress/12200.001.0001>.

And then of course there is the 2021 statement from the Linguistic Society of America on the Scholarly Merit of Open Scholarship in Linguistics: <https://www.linguisticsociety.org/content/statement-scholarly-merit-and-evaluation-open-scholarship-linguistics>.

It seems to me that the authors have not done their due diligence in terms of understanding the state of literature in the field, especially in light of their claims that there is no such literature.

That is the first major weakness of this paper. Misrepresenting the current state of the field makes it very easy for readers to dismiss the topic altogether. Papers urging the research culture of the entire field forward need to carefully and accurately describe the problem at hand, so that readers can recognize the problem for what it is. This in turn makes it harder for readers to ignore the problem, and encourages them to think carefully about how they can help solve it in their own work. My own review is a case in point: I don't recognize the world of linguistics as described in this paper – that is, one in which no linguists have really thought seriously about this issue before today – and I am writing a negative review despite the fact that I hold this topic dear.

We would like to begin by thanking the reviewer for their blunt honesty and the time they have taken to provide references and insight regarding both the topic and our manuscript. Without a doubt, their comments have helped us to reflect on the paper's shortcomings and we believe we have revised accordingly in order to address the two main weaknesses highlighted here. First, we have incorporated all of the suggested references – Berez-Kroeker et al. (2018) and Berez-Kroeker, McDonnell, Koller, and Collister (2022), particularly, but

also Zwaan, Etz, Lucas, and Donnellan (2018) and Gawne and Styles (2022), among others – throughout the revised manuscript. Importantly, we wish to express that we have given careful thought to the assertion that we have misrepresented the current state of the field, which, without a doubt, was not our intention, and have kept this concern present as we have added these key references. In an effort to save space and time, we will not copy/paste all of the changed text in the revised manuscript here, but note the inclusion of the aforementioned references (and others) in paragraphs 1, 2, and 4 of the introduction and paragraphs 2, 4, and 6 of the open data section where we have found them to be particularly relevant.

Additionally, we would like to use a bit of space to address the following comment made by the reviewer: *It seems to me that the authors have not done their due diligence in terms of understanding the state of literature in the field, especially in light of their claims that there is no such literature.* While we disagree with the assertion that we have not done our due diligence, we humbly note that we never make the claim that there is no literature on open science in the field of linguistics. We do, in fact, cite a great deal of it. In the revised manuscript we have made it more clear that our work does not intend to present itself as being singular in this area, but rather as a complement to current efforts. There is no doubt that we did not cite *all* the relevant literature in the original version of the manuscript—certainly this is still the case in the revised version—but we maintain, then and now, that “few, if any, researchers have had explicit instruction on the practices of open science as part of their professional training” and that “current senior researchers were not trained in these innovative methodologies”, which has led to a situation in which “[...] all parties are forced to learn open science on their own, often without institutional support”. This is the reality of our research team and that of most of the researchers with whom we are acquainted. We know that this is certainly not the experience of everybody. Our intent is merely to address the issue as we understand it head on and to provide a helpful resource that can complement the literature that is currently available. In our humble view, the more books, articles, websites, blogs, voices, etc., pushing for open science practices in the field, the better. We do not claim to be pioneers in any way, but we are certain that our contribution has the potential to reach some individuals that are not familiar with any of the extant literature, and that others can benefit from revisiting familiar issues from our perspective. We view our contribution as additive in nature.

[RC 1.2.] The second weakness is that the eight steps outlined here are handled only superficially, and in some cases, in a way that can be potentially harmful. Each section could be a paper unto itself (and interesting special issue idea, perhaps?). Below I highlight a few of the major oversights in the sections.

Positionality statements. The section on positionality statements does not address the counterpoints that have been made regarding the extra burden that required statements put onto BIPOC authors. See, e.g.:

Massoud, Mark Fathi. 2022. The price of positionality: Assessing the benefits and burdens of self-identification in research methods. Journal of Law & Society 1-23.

Roberts, Steven O., et al. 2020. Racial inequality in psychological research: Trends of the past and recommendations for the future. Perspectives on Psychological Science 15: <https://doi.org/10.1177/1745691620927709>.

Open data. The section on open data, while acknowledging that privacy concerns should be respected, does not mention any discussion of the concerns of Indigenous peoples regarding colonial practices in linguistics and especially around who may own, and therefore share, language data (see eg Holton, Leonard & Pulsifer, chapter 4 of OHLDM and that chapter’s very helpful list of references).

Preregistration and registered reports. The section on registered reports does not take into consideration

very real pressures of the systems of academic reward, which bias university researchers against any additional steps in the research process that increase the time to publication (that is, the authors overlook the need for systemic change in academia to encourage open science – see Alperin et al, chapter 13 of OHLDM).

To address the second weakness pointed out by the reviewer, particularly with regard to our treatment of positionality statements, open data, preregistration and registered reports, we have considered the specific issues pointed out above, and, where relevant, expanded our discussion, reorganized, and included critical, missing references (e.g., Holton, Leonard, & Pulisifer, 2022; Leonard, 2021; Massoud, 2022; Mufwene, 2020; Roberts, Bareket-Shavit, Dollins, Goldie, & Mortenson, 2020; Singh, Killen, & Smetana, 2023; Tsikewa, 2021, among others).

Regarding positionality statements, we have heavily revised and reorganized this section. We have focused on including a discussion of the counterpoints of the inclusion of positionality statements particularly regarding marginalized individuals. As this entire section has been reorganized, we direct the reviewer to pages 7-10 of the revised manuscript.

Regarding open data, we have expanded our discussion regarding marginalized communities. We highlight here some of the changes, but also encourage the reviewer to consider the entirety of the revised section in context (see pages 10-14).

Having stated all the above, it is necessary to recognize that the field of linguistics faces unique challenges with regard to open data. At the heart of these challenges are ethical concerns that must be considered with care. First and foremost, the privacy and consent of participants must be safeguarded. Linguistic data often include personal information, which can be especially difficult to anonymize. While on the surface written and behavioral data may not appear to pose as many issues as audio and video recordings, which constitute a large portion of linguistic research materials, it is imperative that one consider the sources from which all types of data are derived. As expressed by Holton et al. (2022), if we haphazardly take language to represent trivial data points and lose focus on the individual embedded within a community, we are doomed to “dehumanize and decontextualize” it (p. 50). This is particularly true when working with minority languages and/or marginalized communities. In cases such as these, the researcher must be held accountable not only for the anonymization of participant information, but also for respecting and upholding the specific goals and restrictions put forth by the community. This includes, but is not limited to, the use, access and storage of all collected data. In sum, careful consideration of the priorities of the researcher and the researched, which often do not align, is paramount (for more detailed views, see Adetula, Forscher, Basnight-Brown, Azouaghe, & IJzerman, 2022; Holton et al., 2022; Hudley, Mallinson, & Bucholtz, 2020; Leonard, 2021; Mufwene, 2020; Singh et al., 2023; Tsikewa, 2021, among others). In addition, generative artificial intelligence technologies, such as Large Language Models, are burgeoning. These technologies will certainly pose currently unknown challenges in the near future that necessitate additional steps to secure the protection of sensitive data against misuse, and uphold the conditions of use put forth by the stakeholders of marginalized communities.

To summarize, open data is important because it facilitates transparency, rigor, reproducibility, replication, accumulation of knowledge, and, importantly, it makes participating in the scientific endeavor more inclusive. According to some accounts, linguistics, in general, does not engage in open science practices, including sharing data (See Bochynska et al., 2023), though others characterize its participation in different terms. For instance, as stated in Berez-Kroeker et al. (2018) “Practitioners in different subfields ‘do transparency’ differently, and these practices could serve as models for an eventual amalgamated standard” (p. 9). While linguistics does face legitimate, field-specific challenges related to non-WEIRD communities, ultimately, the benefits of open data outnumber many of these challenges. Researchers should take the stance to share what is reasonable and ethically responsible all the while holding at the forefront the priorities of the individuals from which the data are derived, especially regarding marginalized communities.

Finally, we have also included discussion of incentive structure in academia and its effects on HARKing, p-hacking and early career researchers.

Researchers have pointed to questionable research practices (QRPs), such as p-hacking—knowingly manipulating an analysis until a significant p-value is obtained (See Head, Holman, Lanfear, Kahn, & Jennions, 2015)—and HARKing—hypothesizing after the results are known (See Murphy & Aguinis, 2019)—, along with small sample sizes, poor theory, lack of transparency, misguided incentive structure in academia, etc., as factors that ultimately led to the replication crisis, though it is likely that many factors are/were simultaneously at play. To wit, many of the aforementioned QRPs may be an unfortunate consequence of misaligned incentive structures in academia, where publication is the universal currency. The pervasive pressure to publish likely leads many researchers to focus on quantity over quality. Couple this with the difficulty of publishing negative or null results, and the result is a research landscape in which many fields suffer from publication bias and there is little to no incentive to prioritize time consuming open science practices. Taking this into account, it is not hard to understand why QRPs might be perceived as a necessary evil by some.

Again, we only highlight some of the changes here, but suggest a reread of the relevant sections in the revised manuscript.

2. Reviewer #2

[RC 2.1.] The argumentation and description of methods is generally very abstract. If I understand the title “opening open science to all” correctly, the authors should include more concrete examples and instructions so that readers who are (still) unfamiliar with the practices described can implement new methods. I added concrete suggestions to do so below; this is necessary for the manuscript to fulfill its promise of opening open science and to become an important resource in research and teaching.

We thank the reviewer for their thoughtful insight. We took a close look at all of the issues raised and have made substantial revisions to the manuscript. Concretely, we have included many more examples and tried to be more inclusive by including in our discussions more of the distinct subfields of linguistics. Furthermore, we have taken the reviewer’s advice and included 3 tables related to preprint servers, registered reports and preregistration, along with 2 new figures. We feel that the manuscript is now greatly improved and we appreciate the time and effort spent giving us such thorough feedback. Below we address all of the concerns raised, and, where relevant, copy/paste prose, figures, and tables.

[RC 2.2.] ERC: should be spelled out here for clarity, as it is only explained later in the introduction.

This acronym has been spelled out in the abstract of the revised manuscript.

[RC 2.3.] “Researchers have pointed to questionable research practices (QRPs), such as p-hacking and HARKing” → HARKing (maybe also p-hacking) needs to be explained briefly here, especially given the goal of the paper.

In the revised manuscript we have defined both terms, included references, and elaborated a bit more on the relationship between QRPs and incentive structures in academia.

Researchers have pointed to questionable research practices (QRPs), such as p-hacking—knowingly manipulating an analysis until a significant p-value is obtained (See Head et al., 2015)—and HARKing—hypothesizing after the results are known (See Murphy & Aguinis, 2019)—, along with small sample sizes, poor theory, lack of transparency, misguided incentive structure in academia, etc., as factors that ultimately led to the replication crisis, though it is likely that many factors are/were simultaneously at play. To wit, many of the aforementioned QRPs may be an unfortunate consequence of misaligned incentive structures in academia, where publication is the universal currency. The pervasive pressure to publish likely leads many researchers to focus on quantity over quality. Couple this with the difficulty of publishing negative or null results, and the result is a research landscape in which many fields suffer from publication bias and there is little to no incentive to prioritize time consuming open science practices. Taking this into account, it is not hard to understand why QRPs might be perceived as a necessary evil by some.

[RC 2.4.] “It necessitates that researchers implement new techniques with limited pedagogical resources and embrace alternative methods of disseminating their research, all of which constitutes a steep learning curve.” → Does it always, though? For instance, in the field of language typology, there are a number of qualitative studies that are based on a sample and annotations using reference grammars. Those studies do not include any code, making them transparent simply means providing a spreadsheet with the languages and the respective annotations and sources that the authors should have in some

form anyway. Can we really speak of “innovative methodologies”? All this goes to say that, yes, some formats (e.g. using OSF) may involve a learning curve, but this is not necessarily so. It may also be about vulnerability, i.e. making all details of the study and analysis public and thus subject to potential criticism, whereas before, those could be left partially “hidden” and “protected”. I suggest that the authors slightly rephrase this paragraph.

We appreciate the reviewer’s thoughtful feedback. We agree that the challenges associated with increasing transparency vary across different fields and methodologies. In response, we have rephrased the paragraph to acknowledge that while some researchers may face a learning curve or feelings of vulnerability when adopting new transparency practices, this is not universally the case. The revised paragraph now reads as follows:

To wit, it often requires learning new skills, thoughtful planning, as well as an openness and willingness to share materials, code, and data. Many researchers need to implement new techniques with limited pedagogical resources and embrace alternative methods of disseminating their research, all of which can constitute a steep learning curve. That being said, what engaging in open science ultimately entails is sure to be field-specific and vary accordingly. In some disciplines, for instance, it may only involve a few of the practices we outline in the present work without the need for innovative methodologies.

[RC 2.5.] **“To this end, we identify three areas, stance, workflow, and dissemination, in which linguists can engage in open science.” → rephrase: “... we identify the following three areas of stance, workflow and dissemination, ...**

This change has been included in the revised manuscript.

[RC 2.6.] **Figure 1 → I am not sure open data is clearly only related to stance; I would also view providing the datasets in a useful form as part of the workflow. Could this point be put in between the two areas?**

The reviewer makes a good point. Open data, in theory, could fit in all three areas depending on what exactly one is referring to, i.e., the stance to make data open or not, how making data open is incorporated into one’s workflow, and how one goes about actually disseminating the data. We have overhauled the figure and include it here for convenience (see below).

[RC 2.7.] **Could you give a concrete example of a positionality statement early on? A concrete example is valuable to make this more understandable for researchers not familiar with this practice.**

In the revised manuscript this section has been reorganized to reflect changes suggested by all three reviewers. We follow the reviewers suggestion and now provide an example of a positionality statement in prose in the first paragraph, with links to more examples at the end of the section.

[RC 2.8.] **“Moreover, Bucholtz et al. (2023) note that considering a researcher’s positionality may be especially important in linguistics, “[...] which relies on racially minoritized communities as sources of data yet lack adequate (if any) representation of those communities among faculty researchers” (p. 2).” → Does this quote really apply to linguistics in general? I can see how it is very relevant for certain areas of linguistics, but there is much linguistic research on major languages, where this point does not seem to apply. Please rephrase / explain.**

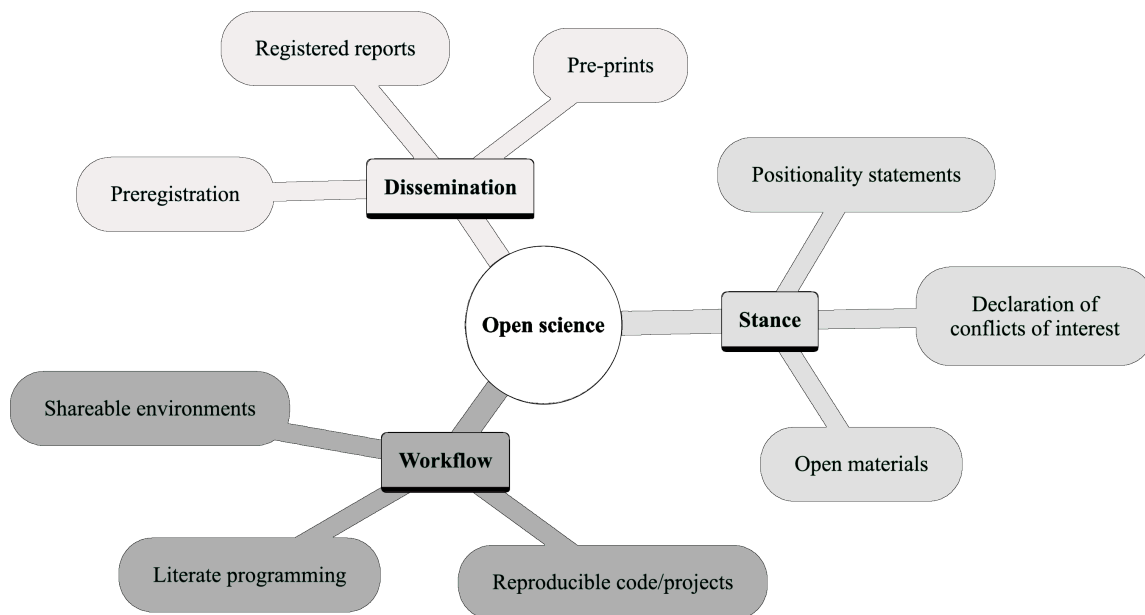


Figure 1: Some open science practices amenable to research in linguistics as they pertain to one's stance, workflow, and the dissemination of research products.

The reviewer's point is duly noted. In the revised manuscript we have rephrased this sentence to make it more clear that we are referring to work on minority languages. The sentence in question now reads as follows:

Moreover, Bucholtz et al. (2023) note that considering a researcher's positionality may be especially important in linguistic research on certain language communities, such as indigenous communities, "[...] which relies on racially minoritized communities as sources of data yet lack adequate (if any) representation of those communities among faculty researchers" (p. 2).

[RC 2.9.] "Many researchers support and advocate for the inclusion of positionality statements in their research publications (e.g., Bucholtz et al., 2023; Jafar, 2018; Steltenpohl et al., 2022)." -> Please specify the research areas here, also those mentioned as arguing against positionality statements.

This section has been heavily edited and reorganized in the revised manuscript. This particular sentence no longer appears as such. That being said, the revised version now expands the discussion around positionality to include more research areas and dedicates more space to discussing the arguments against them.

[RC 2.10.] "Bochynska et al. (2023) surveyed open and transparent practices in linguistics and found that only 10% of the articles sampled included statements of conflict of interest, and, among those 10%, none declared any conflicts (See also Cristea & Ioannidis, 2018; Hardwicke et al., 2022, 2020)." -> Out of curiosity, can you give an example of a statement of conflict of interest from linguistics? Again, concrete examples would help readers less aware of / familiar with these issues understand what those statements can look like.

We thank the reviewer for this welcomed suggestion. We have added references and links to examples in linguistics, psychiatry, psychology, and the social sciences more broadly. We include the revised text for convenience.

Bochynska et al. (2023) surveyed open and transparent practices in linguistics and found that only 10% of the articles sampled included statements of conflict of interest, and, among those 10%, none declared any conflicts. For a clear example of what a declaration of conflicts of interest can entail in linguistics, the interested reader is directed to Bochynska et al. (2023). Of particular value are the *Competing interests* section and the coding form available at <https://escholarship.org/uc/item/6m62j7p6#main> and <https://osf.io/ehyx3>, respectively. Additionally, Cristea and Ioannidis (2018), Hardwicke et al. (2022), and Hardwicke et al. (2020) represent illustrative examples in psychiatry, psychology, and the social sciences more broadly.

[RC 2.11.] **“While positionality statements, due to their reflexive nature, may encompass larger pieces of writing, they can also take the form of short paragraphs that illustrate a few personal characteristics deemed relevant for the particular research endeavor.” → This whole paragraph should be moved to the beginning of the section.**

This section was heavily edited and this particular part, in its revised form, is not at the beginning of the section.

[RC 2.12.] **“Gabriela is a white immigrant cis-gender woman from Romania whose research focuses on how non-native speakers are ideologically framed as linguistically deficient in comparison to native speakers who are characterized by their linguistic authority and expertise.” → OK, this is what I wanted to see early on in this section. Please add some context: What was the title / topic of the publication that this statement belongs to? Can you give a concrete reference? Or is this a made-up example? Please clarify.**

This particular example, which now appears at the beginning of the section, is for one of the authors of the present work. For the sake of anonymity during the peer review process, more details will be included at a later stage.

[RC 2.13.] **“For examples of positionality statements in linguistic research, the interested reader is directed to Bochynska et al. (2023) and Weissler et al. (2023).” → Great pointer to ressources, but if the purpose of this paper really is to make open science practices more accessible to linguists who are not (yet) familiar with them, a concrete example in the beginning of this section is important. It would also help to make this proposal of including a positionality statement more concrete: In which areas of linguistics is it helpful? In which areas is it necessary? In which areas may it be less relevant? For which methods may this be particularly helpful / important? While I sympathize with the arguments made in this section in general, I think they are too abstract in order to be immediately helpful for a concrete implementation. Please make the argumentation a bit more concrete and geared towards certain / different areas of linguistic research so that the reader has more practical advice on how to include such statements in their future research.**

We thank the reviewer for the aforementioned suggestions. They have guided our rewriting/reorganizing of this section, which we believe is now much more informative and useful. As opposed to copy/pasting the entire section here, we encourage the reviewer to consider the revised section in its entirety (approx. pages 7-10).

[RC 2.14.] **“In academic research, statements such as “data available upon request” are commonplace.” → This may be so, but the statement would clearly look better with references or some data to back it up. I do not mean that the authors should single out studies as negative examples, but maybe it would be possible to mention journals and years of publication where this happens. This makes it somewhat more tangible and related to linguistic research. Please rephrase.**

We have added a reference to the aforementioned sentence (See Hardwicke & Ioannidis, 2018), which contains enlightening statistics from the field of psychology.

[RC 2.15.] **“Recent efforts have attempted to encourage researchers to make linguistic data open and accessible via servers (e.g., the IRIS database).” → Given the goal of this paper, please add more explanation about the IRIS database here.**

“open science badges” → Can you explain briefly what this is?

In the revised manuscript we have expanded upon our discussion of the IRIS database and open science badges. The relevant text is included below.

Recent efforts have attempted to encourage researchers to make linguistic data open and accessible via servers. An illustrative example is the IRIS database (<https://www.iris-database.org>), a language sciences digital repository that is freely accessible and permits the up- and downloading of research instruments and materials. Additional efforts include open science badges—visual symbols offered by some journals (e.g., *Language Learning*, *Language and Speech*) on published articles. These badges are awarded to researchers for adhering to certain open science principles, such sharing code and data or preregistering a study. In arguably more extreme cases, other journals have made data sharing a requirement for publication (e.g., *Applied Psycholinguistics*).

[RC 2.16.] **“Though researchers may understandably hesitate to share their data, we believe understanding the benefits of open data can help alleviate any concerns.” → This point needs to be elaborated more in that there may be very different motivations for preferring not to share data. Corpus data may contain sensitive personal information (e.g. video, audio) and speakers may not consent to making it public. Authors of the study may not have the rights to publish the data that their study is based. In this case, it may be possible to make processed data publicly available, but not the raw data (i.e. data can mean different things, and should be distinguished in more detail here). In other cases, it may simply be the fear of making all data and annotation choices criticisable by making everything publicly available. Those are very different perspectives that call for very different solutions / approaches. This needs to be discussed and differentiated in more detail.**

We thank the reviewer for these useful insights. In the revised manuscript we have reorganized this section and rewritten many paragraphs to better reflect these concerns. In addition, we have addressed how certain choices become criticisable upon being made public and included a relevant reference (See Stieglitz et al., 2020). Given the large amount of changes, we do not copy/paste any text here.

[RC 2.17.] **“It affords third parties the opportunity to scrutinize original findings, which promotes reproducibility and reduces errors, such as those related to statistical analyses and reporting of outcomes [TIMO].” → What is [TIMO]?**

We apologize for this oversight. TIMO was a ‘note to self’ to include a reference that was not deleted before submitting the initial version of the manuscript. This has been corrected in the revised version.

[RC 2.18.] **“Revisiting old data sets using innovative techniques can support or contradict past narrative conclusions (e.g., Casillas, 2021).” → It would be useful to elaborate this in a few sentences and present the Casillas (2021) study in terms of revisiting old data sets and drawing (new?) conclusions.**

The revised manuscript now includes a brief discussion of this study and its relevance for open data, which is included below.

Revisiting old data sets using innovative techniques can support or contradict past narrative conclusions. For instance, using meta-analytic techniques, Casillas (2021) reexamined extant research regarding ‘compromise categories’ in early bilinguals. This line of research posits that bilingual individuals produce speech sounds intermediate to those produced by monolingual speakers of either language. By systematically reevaluating prior data and incorporating new acoustic analyses of coronal stops from early Spanish-English bilinguals, Casillas (2021) suggested that the cumulative evidence for ‘compromise’ stop categories was negligible. In lieu of intermediate phonetic categories, the study proposed early bilinguals can exhibit performance mismatches resulting from dynamic interlingual interactions. This reanalysis contradicted earlier assumptions about bilingual phonology and provided in-depth scrutiny of statistical power and evidence accumulation in bilingualism research.

[RC 2.19.] **“Open data are particularly important for the field of linguistics, for all of the aforementioned reasons, and also because linguistics is Western, Educated, Industrialized, Rich, and Democratic (WEIRD, Bochynska et al., 2023).” → I do not disagree, in principle, but there may be one caveat that needs to be mentioned: Bochynska et al. (2023) evaluated publications written in English. This seems slightly self-selecting to me, as it probably excludes relevant linguistic research traditions whose main language of publication is not English (e.g. Chinese linguistics). How sure can we then be that linguistics as a whole is WEIRD? I suggest the authors tone down this statement a bit, as it seems too strong in its current form.**

We agree with the reviewer’s assessment and have toned down the message accordingly. In addition, we have added more references that highlight this assertion about the field. The revised text is included below.

Open data are particularly important for the field of linguistics, for all of the aforementioned reasons, and also because some linguists have described the state of the field, as far as English-language publications are concerned, as being Western, Educated, Industrialized, Rich, and Democratic (WEIRD, see Bochynska et al., 2023; Faytak et al., 2024; Nagle, Baese-Berk, Amengual, & Casillas, 2024). That is to say, the majority of linguistic research appears to be concentrated on specific languages, mainly Indo-Germanic, in overrepresented communities, by privileged scholars.

[RC 2.20.] **“Making linguistic data accessible to all researchers promotes participation in and with underrepresented communities. Furthermore, it can increase the study of diverse and underreported languages, which fosters a more inclusive and comprehensive understanding of the global linguistic landscape.” → Please explain why and how accessible linguistic data promotes the researchers’ participation with the (potentially underrepresented) speaker communities? While I am sympathetic to this argument, it needs to be spelled out more and made more concrete in order to be meaningful.**

We have tried to make this point more clear in the revised manuscript. We include the relevant paragraph here for convenience.

Open data are particularly important for the field of linguistics, for all of the aforementioned reasons, and also because some linguists have described the state of the field, as far as English-language publications are concerned, as being Western, Educated, Industrialized, Rich, and Democratic (WEIRD, see Bochynska et al., 2023; Faytak et al., 2024; Nagle et al., 2024). That is to say, the majority of linguistic research appears to be concentrated on specific languages, mainly Indo-Germanic, in overrepresented communities, by privileged scholars. Making linguistic data accessible to all researchers can promote participation *in* and *with* underrepresented communities. Furthermore, it can increase the study of diverse and underreported languages by affording more researchers the opportunity to interact and learn from data that would otherwise not be available to them, which, in turn, can foster a more inclusive and comprehensive understanding of the global linguistic landscape.

[RC 2.21.] **“Having stated all the above, it is necessary to recognize that the field of linguistics faces unique challenges with regard to open data.” → Are these necessarily unique challenges? I do not think it is important to point this out (and I am not sure if linguistics really faces unique challenges that different from other related disciplines). What would be more important is to go into more detail about which areas of linguistics face which types of challenges, since there are so many different types of data that we as linguists work with (experimental data, corpus data, data from grammars, data from linguistic databases, written, spoken, signed data, historical data, and each of these types could be distinguished further). This should be discussed in much more detail and made more concrete so that the readers can see their approaches / types of data represented in this discussion.**

We again thank the reviewer for this helpful comment. This section has been reworked to address the challenges researchers face in the field of linguistics, and provide a more nuanced view regarding marginalized communities. We refer the reviewer to pages 11-16 (approx.) of the revised manuscript.

[RC 2.22.] **“The privacy and consent of participants must be safeguarded...” → OK, this is the kind of information that I was waiting for earlier; consider moving this paragraph up. Also, not all linguistic studies face ethical issues: What about studies that work with databases / corpora that have been compiled by third parties, which are publicly accessible and which the authors simply use as is. They are far removed from any speakers in this case and had no influence on any decision of data collection and processing that went into compiling the database. Or what about historical linguistics? The role that participants or informants play in different types of linguistic studies deserves a more fine-grained discussion.**

Following the reviewer’s suggestion, this section was reorganized to address their concerns. We refer the reviewer to pages 11-16 (approx.) of the revised manuscript.

[RC 2.23.] **“Moreover, the advent of generative artificial intelligence technologies, such as Large Language Models, may pose unknown challenges in the near future that necessitate additional steps to secure the protection of sensitive data against misuse.” → Please explain in more detail: Is the idea that LLMs use data compiled by linguists for linguistic research? Or does the potential of misuse lie in the linguistic research with LLMs itself? If the authors wish to mention this, they should discuss in more detail what exactly the risks are. Otherwise, this sentence could be omitted.**

We thank the reviewer for recommending clarity regarding this point. We are referring to the fact that LLMs

are new and we don't fully understand how they will be used in the future, by both academics and the general public. Given the fact that they operate using/are trained by large amounts of text data, the availability of open data/corpora could lead to potentially sensitive data (e.g., data derived from marginalized communities) being used in ways that do not adhere to the original agreement of informed consent. We have clarified this point in the revised manuscript and include the relevant text below.

In addition, generative artificial intelligence technologies, such as Large Language Models, are burgeoning. These technologies will certainly pose currently unknown challenges in the near future and may necessitate additional steps to secure the protection of sensitive data against misuse, particularly regarding adherence to the original agreement of informed consent, and, importantly, in upholding the conditions of use put forth by the stakeholders in marginalized communities.

[RC 2.24.] **“When primary data, such as audio or video files, cannot be shared, derived data in the form of tabular files can take its place...” -> This paragraph is very important and constructive. Different types of data (primary data vs. processed data and annotations) should be introduced at the beginning of this section, though, and the entire section should be more precise about what types of data particular arguments refer to.**

As mentioned above, this section has been rewritten/reorganized to reflect suggestions by all reviewers. We refer the reviewer to pages 11-16 (approx.) of the revised manuscript.

[RC 2.25.] **Prolific -> add a link, e.g. in a footnote.**

A link to the prolific website has been included in the revised manuscript.

[RC 2.26.] **“In more uncommon cases in which institutional policies do not permit the sharing of derived data sets, synthetic data containing the same statistical properties can be generated and shared freely (See Quintana, 2020).” -> Please elaborate how this would work.**

In the revised manuscript we have included more information regarding how the process works and provided links to code and an example tutorial.

In more uncommon cases in which institutional policies do not permit the sharing of derived data sets, synthetic data containing the same statistical properties can be generated and shared freely (See Quintana, 2020). In short, the method consists of capturing the statistical properties of the original data set and using them to simulate new data that preserve the relationships between the variables of interest. A tutorial on the method described in Quintana (2020) is freely available on github (<https://github.com/elifesciences-publications/synthpop-primer>), and an online RStudio instance can be accessed at <https://mybinder.org/v2/gh/dsquintana/synthpop-primer/master?urlpath=rstudio>. All relevant materials are available on the OSF: <https://osf.io/z524n/>.

[RC 2.27.] **“A substantial hurdle that cannot be overlooked revolves around the fact that researchers must learn to use new technologies to participate in open, transparent research. Making data open and accessible is not as simple as merely uploading a data file.” -> I disagree in that it depends on the linguistic subdiscipline. In large-scale typological studies that are qualitative in nature and use no statistics but are**

based on a language sample, making research transparent simply consists of adding a supplementary spread sheet file with the names of the languages, all relevant annotations / examples and their sources. Impressionistically, even this is still not done in all studies, so improving on transparency in this case only involves adding a supplementary file to the publication. Also, if the researcher does not provide such supplementary files on their own, many journals offer to host supplementary materials on the journal website together with the publication. This may not be the best solution, but it is a solution where the researcher really does not need to know or consider platforms to host data. Please clarify this.

The reviewer's disagreement regarding this issue is duly noted. We believe we failed to fully describe and contextualize the difficulties involved with making data open and accessible. In the revised manuscript we have further elaborated, providing examples and templates. For the sake of completeness, we describe the issue here. The logic behind our argument revolves around the fact that simply providing access to data (or a data file) is not sufficient. To elaborate using the reviewer's example, we will consider the case in which a supplementary spread sheet file with the names of languages represent the data in question. In this case, uploading the file, be it on an open, independent server or on the journal's website, does not provide the relevant context for a non-expert to use the data. This is the case because, when publicly sharing data, it is ideal that one also provide three distinct levels of documentation so that an independent researcher has the proper context to use data. It is helpful to have in mind a researcher that may not be completely familiar with the norms of a particular subfield. This part is key because the objective is that any individual be able to use the data. The levels of documentation are *project-level* (i.e., a project summary document), *data-level* (i.e., a README file explaining the data set), and *variable-level* (i.e., a data dictionary). The different levels will be more or less relevant depending on the subfield. In our example, it could theoretically be critical for an independent researcher to understand how and why the data are relevant to the broader goals of the project (project-level), as well as how and why the contents of the data file (in our case the languages) came to be included and any relevant relationships or hierarchies among them (data-level, e.g., language families, varieties of a single language, etc). In this particular case, a data dictionary (variable-level) could also be relevant to understand how the languages are listed, i.e., whether or not abbreviations are used, or if the names are interpretable to speakers of other languages. If it wasn't obvious, this particular example is not in the area of our (the authors) collective expertise. Nonetheless, we believe this makes the example even more relevant, as the topics we have covered represent aspects of our ignorance in this area that we believe we may need to know in order to make full use of a data set of this nature. Importantly, the type of thoroughness we outline here is virtually non-existent in most data sets publicly available. For this reason we also have included links to templates any researcher could use and adapt for their specific purposes. In the revised manuscript there relevant text reads as follows:

Another substantial hurdle that cannot be overlooked revolves around the fact that researchers must learn to use new technologies to participate in open, transparent research. Making data open *and* accessible is not as simple as merely uploading a data file. Ideally, researchers should include relevant information to contextualize the data set at the project-level (i.e., a project-summary document), the data-level (i.e., a README file explaining the data set), and the variable-level (i.e., a data dictionary) (Lewis, 2024). The inclusion of resources at these three levels is the optimal way for authors to provide the necessary context for an independent researcher to access and utilize the data set. Unfortunately, most publicly available data do not adhere to this standard. For this reason, we direct the interested reader to templates provided in Lewis (2024) for documentation at the project-level (<https://osf.io/q6g8d>), data-level (<https://osf.io/d3pum>), data-level (<https://osf.io/tk4cb>), and variable-level (<https://osf.io/ynqcu>).

[RC 2.28.] “Free repositories designed for the purpose of sharing research materials, such as the Open Science Framework (Foster & Deardorff, 2017), github, etc., are preferable and can be accessed simply by sharing a link.” → Here, it would actually be helpful to spell this out a bit more and mention e.g. zenodo as another option. For readers less familiar with these options, it would be valuable to add a table with different platforms / systems and provide information about long-term support, version control, possibility of DOI, anonymous links (e.g. what OSF offers and what is very useful for sharing data and code during the revision stage of a study), etc.

We thank the reviewer for this idea. In the revised manuscript we have included a table summarizing some of the available resources. We include the table here for convenience.

Platform	Long-term Support	Version Control	DOI Assignment	Anonymous Sharing	Key Features
Open Science Framework (OSF)	+	+	+	+	Project management and collaboration
GitHub	+	+	Integrates with Zenodo	+(public repositories)	Project management/collaboration, ideal for coding
GitLab	+	+	Integrates with Zenodo	Limited	Project management/collaboration, ideal for coding
Bitbucket	+	+	Integrates with Zenodo	Limited	Project management/collaboration, ideal for coding
Zenodo	+	+(via GitHub integration)	+	+	Supports range of file types
Figshare	+	Limited	+	+	Sharing datasets and figures
Box	-	-	-	Limited	Basic file storage and sharing
Google Drive	-	-	-	Limited	Basic file storage and sharing

[RC 2.29.] “While linguistics does face legitimate, field-specific challenges, ultimately the benefits of open data outnumber these challenges, and researchers should take the stance to share what is ethically reasonable.” → My impression when reading this is that the only field-specific challenge mentioned was the point on sensitive speaker data. I generally agree with the statement, but in order to make this statement at the end of the section, the section would need to be much more explicit about the different challenges taking into account the different areas of linguistics.

Another point to consider is the following, although it may not immediately relate to data and more to the inclusion of lesser studied languages and integration of speaker communities in linguistic research:

There are linguistic journals that allow / require an additional abstract of the papers in the target language or another local language besides the usual abstract in English. From personal experience, I know that Linguistics Vanguard offers this as an option (although I am not sure if you can find this information other than submitting a paper). Another journal that makes use of this practice is the Journal of African Languages and Linguistics (just to give some examples from the most recent volume: <https://www.degruyter.com/document/doi/10.1515/jall-2023-2008/html>, <https://www.degruyter.com/document/doi/10.1515/jall-2023-2011/html>)

Again, the reviewer's point is duly noted. We have reorganized and rewritten large portions of this section to include more field-specific challenges. For the sake of convenience we highlight but one here and recommend a reread of the section in its entirety.

Having stated all the above, it is necessary to recognize that linguistics faces a unique set of challenges with its multitude of subfields, each potentially working with a variety of data formats. Due to such diversity, one must determine which aspects of open science are relevant to their data. For example, a neurolinguistic study investigating event related potentials (ERPs) could share raw data for transparency, as well as preprocessed data with the code used to transform the raw data and a corresponding description for facilitation of reanalysis. In another field, the creation of a corpus will benefit from open access and the use of standardized file formats; the analysis of a corpus will benefit from sharing the search queries, the analysis code, and a description of the analysis code. At the heart of these challenges are ethical concerns that must be considered with care.

[RC 2.30.] "As we have seen, reproducibility is now a crucial aspect of any scientific study." -> Where did we see that? Also, I agree that it should be, but reproducibility is not yet a crucial aspect of any scientific study, unfortunately. Please rephrase.

We have rephrased this sentence in the revised manuscript. It now reads as follows:

Having seen the consequences from the reproducibility crisis in other fields, reproducibility must be a crucial aspect of any scientific study.

[RC 2.31.] "At worst, a lack of reproducibility can lead to irreproducible results and wasted resources." -> Is this really the worst consequence? Is it not worse (or equally bad) that we may end up with results that are taken as a given for decades in a given linguistic field, because the original study was not reproducible and never replicated but influential for some reason?

We have rewritten this paragraph. In the revised manuscript it now appears as follows:

In general, reproducibility helps to increase the credibility of research findings and allows other researchers to verify and build on existing work. A lack of reproducibility can lead to findings that cannot be replicated, resulting in wasted resources, and, conceivably, downstream impacts on public health and policy decisions that are often grounded in funded research. For these reasons, among others, transparency in research methods are essential to ensure reproducibility, which includes not only the data collection and analysis methods, but also the code used to conduct the analysis. In linguistics there is increasing awareness of the importance of reproducibility and how a lack thereof could potentially impede advancements in linguistic theory and theories of language acquisition, in addition to having implications for education and language policy decisions based on research findings. As a consequence, many investigators are showing heightened interest in safeguarding the reproducibility of their research.

[RC 2.32.] **“This can have serious implications for public health and policy decisions based on research findings.”**
 -> Is this statement about linguistics or research in general? Please clarify.

“In linguistics there is increasing awareness of the importance of reproducibility, and many researchers are beginning to take steps to improve the reproducibility of their research.” -> too many reproducibilities in one sentence, rephrase.

These concern were addressed in the rewriting of the original paragraph (see response directly above).

[RC 2.33.] **“There are several steps that researchers can take to make their code and projects more reproducible.”**
 -> A large portion of linguistic work is qualitative and does not rely on any code. If the paper is meant to be relevant to all linguistic approaches in general, please be more inclusive. Alternatively, the topic of the paper could be specified more and refer to “quantitative linguistics”?

Again, the reviewer’s point is duly noted. There are some areas of open science that are clearly more applicable to certain subfields of linguistics. As the reviewer points out, this particular section is more relevant for quantitative research. We have rewritten the beginning of this paragraph to highlight this fact. The revised text now reads as follows.

For quantitative research, there are several steps that researchers can take to make their code and projects more reproducible. One approach is to create reports that document the research process by including descriptions of the data, the methods used to analyze the data, and the results.

Notably, other sections of the manuscript are more applicable to qualitative research (e.g., preregistration and registered reports) and highlighted at a later juncture.

[RC 2.34.] • **regarding the sharing of code, dependencies, versions & reproducibility: It may be helpful to explain in more detail why it is important to not only share the code, i.e. that different operating systems can react to code differently, functions can differ in different versions of packages, etc. The authors may consider referring to Roberts et al. (2015), who show that functions from the lme4 R package produce different results when used on different operating systems (althouh fixed since then). Roberts, Seán, James Winters & Keith Chen. 2015. Future tense and economic decisions: Controlling for cultural evolution. PLOS ONE 10(7). e0132145.**

<https://doi.org/10.1371/journal.pone.0132145>.

- The literate coding section is very much focused on R. This may be warranted given that R seems to be the main programming language used for statistical analyses in many (?) linguistic areas, but python is certainly not uncommon either. If the authors go into detail for literate programming in R (e.g. citing the knitr package), they should also do this for python, and potentially mention Julia, which is also, although less commonly, used in linguistics. Here again, a table with programming languages, approaches / formats / packages as well as platforms with virtual environments would be very helpful.
- “This includes updating code and documentation as needed and testing projects on different operating systems to ensure that it can be run in different environments.” → This statement seems to assume that researchers should have access to different OSs, which may not always be the case. Also, it is hardly possible to test for all compatibilities and eventualities on the side of the authors. It is more reasonable to create and publish an, e.g., Docker environment for an R project, so that other researchers can reproduce the original analysis with the exact OS and package versions as in the original study.

The reviewer brings up excellent points. We address them here together, as they relate to a common theme. In the revised manuscript we have taken care to explain *why* it is important to share more than just code and how reproducibility can be affected by system and platform specific issues. With regard to the literate programming section, we have made sure to discuss more of the available languages and platforms as well. In this particular case, we opted not to include a table. Additionally, we rephrased our treatment of updating and testing code (we do not intend to imply that a researcher has the obligation to test their projects on all conceivable platforms). To underscore the usefulness of creating public instances of projects (e.g., via Docker), we have made our project available there as well. We mention this in this section and will provide the working link upon completing the project. This is also true for the OSF and github repositories. For the time being, the anonymous version of the OSF repo is available here: https://osf.io/bsu2q/?view_only=68d1e41b327f4a28a9fcd0fc6537ecaf.

[RC 2.35.] “Linguistic research is multifaceted and spans diverse areas such as corpus analysis, conversation/discourse analysis, experimental research, and more.” → This is a very reduced list of linguistic research areas, please add more areas and discuss in more detail for which areas preregistration makes sense. For instance, can it be used in theoretical linguistics? It is mentioned in the beginning that “we focus on who might want to consider preregistrations”, but this does not become very clear. Please add at least a paragraph about what types of linguistic studies can benefit from preregistrations.

We agree with the reviewer’s assessment of this sentence. We have rewritten it to better represent some of the subfields of linguistics.

Linguistic research is multifaceted and spans diverse areas such as phonetics, phonology, syntax, morphology, sociolinguistics, natural language processing, and conversation/discourse analysis, to name just a few. These areas range from purely theoretical to quantitative and experimental, with many falling somewhere in between.

In addition, we have added more detail throughout this section regarding *who* can/should use preprints and for *what* purpose.

[RC 2.36.] **“Researchers face vital decisions while engaging in research, with inherent flexibility involved in the process of designing and conducting experiments, as well as analyzing the results (Simmons, Nelson, & Simonsohn, 2011).” → Does this mean you only consider experimental linguistics research? This needs to be framed differently, or, the scope of the paper should be reduced to experimental linguistics.**

We thank the reviewer for this helpful comment. We have taken care to be more inclusive in our wording, and, in turn, more faithful to the point being made by Simmons, Nelson, and Simonsohn (2011). The revised text now reads as follows:

Researchers face vital decisions while engaging in research, with inherent flexibility involved in the process of designing and carrying out projects, as well as in the analysis of the data and interpretation of the results (Simmons et al., 2011).

[RC 2.37.] **Coretta et al. 2023: Please explain in a few sentences what the study did and what it showed; it is impossible to understand (and appreciate) that without knowing the study.**

We have included a description of the Coretta et al. (2023) study in order to facilitate understanding/appreciating “researcher degrees of freedom”. The revised text is included below.

This type of flexibility, termed “researcher degrees of freedom”, can have serious down-stream consequences in quantitative research, particularly in linguistics. For instance, Coretta et al. (2023) provided the same speech-production data set to different research teams and asked them to answer the same research question. They found substantial variability in both the acoustic analyses and the analytic strategies, neither of which could be explained by analysts’ prior beliefs, expertise, or the perceived quality of their analyses. Crucially, these decisions, both acoustic and analytic, impacted the teams’ answers to the research question. To provide a simple example, a researcher studying lexical stress could concentrate on distinct acoustic cues typically associated with stress, i.e., pitch, duration, and intensity. Beyond selecting acoustic cues to measure, she must also select a domain for these measurements, such as the mid-point of stressed/unstressed syllables or an average value over the entirety of the syllable. Choices such as these, i.e., the researcher degrees of freedom, can wield significant influence on subsequent outcomes. Preregistration serves the purpose of meticulously documenting these choices *a priori*, thus acting as a deterrent against QRPs, like HARKing or p-hacking (Wicherts et al., 2016).

[RC 2.38.] **This section needs more concrete pointers to specific platforms, journals, spaces for preregistration. Linguists outside of experimental research may want to do this, but are less likely to know how and where exactly they can preregister their study. The abstract motivations are spelled out clearly in this section, but it is not of much concrete help to those linguists who are convinced about the method but have no experience with it. It would be helpful to see an example of how this can be done outside of experimental studies, e.g. a corpus study, a typological study, etc. What about e.g. a theoretical morphological / phonological / syntactic analysis based on information from grammars? Is preregistration even possible or sensible? Please add more concrete details regarding different areas of linguistics so that this section becomes useful to the readers**

We thank the reviewer for this excellent suggestion. In the revised manuscript we have included more resources related to preregistration in distinct subfields of linguistics. Though multiple platforms exist for

preregistering studies, we have opted not to include a table summarizing the options. We made this decision based on the fact that there are two primary platforms used generally, OSF and aspredicted.org, the former much more commonly used than the latter. The remaining platforms (e.g., ClinicalTrials.gov, European Union Clinical Trials Register, PROSPERO, SRCD, etc.) are all field-specific, and, therefore, not relevant for linguistics.

[RC 2.39.] Please give examples of journals / platforms that allow for registered reports.

We again thank the reviewer for a excellent comment that resulted in the inclusion of extremely relevant information. The revised manuscript now includes data and a summary table regarding registered reports in linguistics. The reviewer is directed to pages 25-26 (approx.) of the revised manuscript.

[RC 2.40.] “First, select a pre-print server that aligns with the course of research.” → Please give examples of suitable pre-print servers for different linguistic areas. Again, a table would be very helpful.

At the reviewers request, the revised manuscript now includes examples of suitable pre-print servers, as well as a table summarizing some of the interesting details.

Server	Discipline(s)	Database Size	Year Created	URL
LingBuzz	General Linguistics	+8,000	2006	https://ling.auf.net/lingbuzz
Open Science Framework	Multidisciplinary (includes Linguistics)	+3M	2011	https://osf.io/preprints
PsyArXiv	Psychology, Cognitive Sciences, Psycholinguistics, Linguistics	+30,000	2016	https://osf.io/preprints/psyarxiv
Cogprints	Multidisciplinary (includes Cognitive Sciences and Linguistics)	+4,000	1995	https://web-archive.southampton.ac.uk/cogprints.org/
SocArXiv	Social Sciences (includes Sociolinguistics)	+10,000	2016	https://osf.io/preprints/socarxiv
EdArXiv	Education Research (includes Applied Linguistics)	+1,000	2018	https://osf.io/preprints/edarxiv
Computational Linguistics Open Archive (CLARIN)	Language-Based Research	N/A	2012	https://www.clarin.eu/
ACL Anthology	Computational Linguistics and NLP	+10,000	2004	https://aclanthology.org/
arXiv	Multidisciplinary (includes Computational Linguistics, NLP)	+2,5M	1991	https://arxiv.org/
SciELO Preprints	Research pertinent to Latin America, Spain, Portugal and South Africa	+140,000	1998	https://preprints.scielo.org/index.php/scielo/preprints
HAL (Hyper Articles en Ligne)	Multidisciplinary (includes Language-Specific French Linguistics)	+1M	2001	https://hal.science/

[RC 2.41.] **“We have provided descriptions and relevant examples of these practices to accompany the many guides already available for learning open science (e.g., Crüwell et al., 2018; Lewis, 2020). Crucially, the purpose of this article is to help foster open science in linguistics (FOSIL).” → The current version of this manuscript provided detailed motivations and descriptions, but not sufficient examples and concrete instructions for linguists who are less familiar with these practices. Please add more concrete details, examples, instructions as indicated above. → The FOSIL tutorials should be mentioned much earlier whenever relevant, given that they do include more concrete information.**

Following the reviewer’s suggestions, we have included more concrete examples pertaining to linguistics throughout the revised manuscript, and we have referenced the FOSIL tutorials where relevant. The examples are too many to reproduce again here, but we are confident they are more “visible” in the updated version.

[RC 2.42.] **Some references contain incomplete author lists → all authors should be cited.**

We have carefully scrutinized there reference section and made the appropriate changes.

3. Reviewer #3

[RC 3.1.] Since QRPs appear to be a critical component of the commentary here, I would spend at least a paragraph in the introduction explaining them / the original few papers on them and how this interwove with the time period of the 2015 collaboration to lead to TOP guidelines.

We thank the reviewer for this comment, which was common among the other reviewers as well. In the revised manuscript we have dedicated more space to discussing/explaining QRPs and incentive structure in academia. Particularly relevant for the reviewer's comment is the revised text we include below:

Researchers have pointed to questionable research practices (QRPs), such as p-hacking—knowingly manipulating an analysis until a significant p-value is obtained (See Head et al., 2015)—and HARKing—hypothesizing after the results are known (See Murphy & Aguinis, 2019)—, along with small sample sizes, poor theory, lack of transparency, misguided incentive structure in academia, etc., as factors that ultimately led to the replication crisis, though it is likely that many factors are/were simultaneously at play. It is likely that many of the aforementioned QRPs are an unfortunate consequence of misaligned incentive structures in academia, where publication is the universal currency. The pervasive pressure to publish likely leads many researchers to focus on quantity over quality. Couple this with the difficulty of publishing negative or null results, and the result is a research landscape in which many fields suffer from publication bias and there is little to no incentive to prioritize time consuming open science practices. Taking this into account, it is not hard to understand why QRPs might be perceived as a necessary evil by some.

In the aftermath of the aforementioned crisis, there has been a push for increased transparency and reproducible methodology to help mitigate the effects of QRPs. The clearest example of this is the Transparency and Openness Promotion Guidelines (TOP), which are author guidelines for journals that aim to help evaluate adherence to open science principles (See Nosek et al., 2015, as well as <https://www.cos.io/initiatives/top-guidelines>). The resulting methodological framework and associated techniques have reshaped research methods in Psychology, and, slowly but surely, are making their way into related fields. While many agree that open science practices represent a positive step forward in improving scientific rigor, these practices, by and large, have not been adopted in the field of linguistics (Bochynska et al., 2023). One reason for the slow adoption in linguistics may be related to the fact that engaging in open science is no trivial feat. On the contrary, it often requires learning new skills, thoughtful planning, as well as an openness and willingness to share materials, code, and data. Many researchers need to implement new techniques with limited pedagogical resources and embrace alternative methods of disseminating their research, all of which can constitute a steep learning curve. That being said, what engaging in open science ultimately entails is sure to be field-specific and vary accordingly. In some disciplines, for instance, it may only involve a few of the practices we outline in the present work without the need for innovative methodologies. Nonetheless, given how new open science practices are, it is reasonable to assume that current senior researchers were not trained in these innovative methodologies. As a consequence, many early career researchers (ERC) find themselves at a crossroads in which they are forced to learn open science on their own, often without institutional support. Ironically, there is also a growing expectation that ECRs implement these novel tools in order to be successful in their programs, on the job market, or to advance in their careers.

[RC 3.2.] I love this figure. Given the pervasiveness of TOP guidelines, maybe make the connection between them direct. I think you could comment on how those guidelines are heavily quant focused and ignore the intersection of the researcher and the research (and how positionality statements acknowledge these things).

Though this figure has changed significantly, we thank the reviewer for their praise nonetheless. We have also made the connection with the TOP guidelines more clear and heavily revised the positionality statement section.

[RC 3.3.] When you talk about repositories, mention that there are lists of “trusted repositories” that are better than researcher websites, are more accessible, etc.

In the revised manuscript we now include a table that highlights some of the more common data sharing platforms and highlight some of the key features.

[RC 3.4.] Organization: I would make the big headers each of your sections from the figure and then subheader the points in each.

We thank the reviewer for this excellent idea. We have made this organizational change to the revised manuscript.

[RC 3.5.] Maybe provide a link to tutorials that show people how to do literate programming? I think there are a few good markdown / rmarkdown / quarto ones that may be beneficial for further reading. You could also mention code ocean as options for researchers who don’t want to recreate entire environments but want to test the code.

Oh here’s code ocean ha. Maybe a touch earlier.

We again thank the reviewer for this practical and useful idea. We have provided examples of literate programming, highlighting this very manuscript as an example tutorial. We make the entire project available on the OSF, GitHub, as well as an online server instance via Code Ocean.

[RC 3.6.] I think the section on pre-reg should start with a quite note that they are not only for experimental studies, you can pre-reg ideas, etc. I think there’s been a lot of push back against them because it doesn’t feel like it fits in qual research or more exploratory studies.

We wholeheartedly agree with the reviewer’s observation. In the revised text we have tried to emphasize that pre-registrations are adaptable to many subfields of linguistics by providing more concrete examples throughout this section.

[RC 3.7.] In the pre-print section, this is a good place for TOP as well – many of the guidelines encourage them, so that’s a helpful signal.

It is not entirely clear to us how exactly the reviewer envisions promoting TOP in this section. Which of the guidelines encourage pre-prints? We use the OSF (and corresponding publication, Nosek et al., 2015) as our guide, but do not see a clear connection. We are happy to include this in a future revision of the manuscript once we have a better idea of what exactly the reviewer means.

References

- Adetula, A., Forscher, P. S., Basnight-Brown, D., Azouaghe, S., & IJzerman, H. (2022). Psychology should generalize from—not just to—Africa. *Nature Reviews Psychology*, 1(7), 370–371. <https://doi.org/10.1038/s44159-022-00070-y>
- Berez-Kroeker, A. L., Gawne, L., Kung, S. S., Kelly, B. F., Heston, T., Holton, G., ... Woodbury, A. C. (2018). Reproducible research in linguistics: A position statement on data citation and attribution in our field. *Linguistics*, 56(1), 1–18. <https://doi.org/10.1515/ling-2017-0032>
- Berez-Kroeker, A. L., McDonnell, B., Koller, E., & Collister, L. B. (Eds.). (2022). *The open handbook of linguistic data management*. The MIT Press. <https://doi.org/10.7551/mitpress/12200.001.0001>
- Bochynska, A., Keeble, L., Halfacre, C., Casillas, J. V., Champagne, I.-A., Chen, K., ... Roettger, T. B. (2023). Reproducible research practices and transparency across linguistics. *Glossa Psycholinguistics*, 2(1, 18), 1–36. <https://doi.org/10.5070/G6011239>
- Casillas, J. V. (2021). Interlingual interactions elicit performance mismatches not “compromise” categories in early bilinguals: Evidence from meta-analysis and coronal stops. *Languages*, 6(1), 9. <https://doi.org/10.3390/languages6010009>
- Coretta, S., Casillas, J. V., Roessig, S., Franke, M., Ahn, B., Al-Hoorie, A. H., ... Roettger, T. B. (2023). Multidimensional signals and analytic flexibility: Estimating degrees of freedom in human-speech analyses. *Advances in Methods and Practices in Psychological Science*, 6(3), 1–29. <https://doi.org/10.1177/25152459231162567>
- Cristea, I.-A., & Ioannidis, J. P. A. (2018). Improving disclosure of financial conflicts of interest for research on psychosocial interventions. *JAMA Psychiatry*, 75(6), 541–542. <https://doi.org/10.1001/jamapsychiatry.2018.0382>
- Faytak, M., Kadavá, Š., Xu, C., Özsoy, O., Akumbu, PiusW., Cardoso, A., ... Roettger, T. B. (2024). *Big team science for language science: Opportunities and challenges*. Open Science Framework. Retrieved from osf.io/3pkj6
- Gawne, L., & Styles, S. (2022). Situating linguistics in the social science data movement. In A. L. Berez-Kroeker, B. McDonnell, E. Koller, & L. B. Collister (Eds.), *The open handbook of linguistic data management* (pp. 9–25). The MIT Press. <https://doi.org/10.7551/mitpress/12200.003.0006>
- Hardwicke, T. E., & Ioannidis, J. P. (2018). Populating the data ark: An attempt to retrieve, preserve, and liberate data from the most highly-cited psychology and psychiatry articles. *PloS One*, 13(8), 1–12. <https://doi.org/10.1371/journal.pone.0201856>
- Hardwicke, T. E., Thibault, R. T., Kosie, J. E., Wallach, J. D., Kidwell, M. C., & Ioannidis, J. P. A. (2022). Estimating the prevalence of transparency and reproducibility-related research practices in psychology (2014–2017). *Perspectives on Psychological Science*, 17(1), 239–251. <https://doi.org/10.1177/1745691620979806>
- Hardwicke, T. E., Wallach, J. D., Kidwell, M. C., Bendixen, T., Crüwell, S., & Ioannidis, J. P. A. (2020). An empirical assessment of transparency and reproducibility-related research practices in the social sciences (2014–2017). *Royal Society Open Science*, 7(2), 1–10. <https://doi.org/10.1098/rsos.190806>
- Head, M. L., Holman, L., Lanfear, R., Kahn, A. T., & Jennions, M. D. (2015). The extent and consequences of p-hacking in science. *PLoS Biology*, 13(3), 1–15. <https://doi.org/10.1371/journal.pbio.1002106>
- Holton, G., Leonard, W. Y., & Pulisifer, P. L. (2022). Indigenous peoples, ethics, and linguistic data. In A. L. Berez-Kroeker, B. McDonnell, E. Koller, & L. B. Collister (Eds.), *The open handbook of linguistic data*

- management* (pp. 51–60). The MIT Press. <https://doi.org/10.7551/mitpress/12200.003.0008>
- Hudley, A. H. C., Mallinson, C., & Bucholtz, M. (2020). Toward racial justice in linguistics: Interdisciplinary insights into theorizing race in the discipline and diversifying the profession. *Language*, 96(4), e200–e235. <https://doi.org/10.1353/lan.2020.0074>
- Leonard, W. Y. (2021). Centering indigenous ways of knowing in collaborative language work. In L. Crowshow, I. Genée, M. Peddle, J. Smith, & C. Snoek (Eds.), *Sustaining indigenous languages: Connecting communities, teachers, and scholars* (pp. 21–34). Athabasca University Press.
- Lewis, C. (2024). *Data management in large-scale education research*. CRC Press.
- Massoud, M. F. (2022). The price of positionality: Assessing the benefits and burdens of self-identification in research methods. *Journal of Law and Society*, 49, S64–S86. <https://doi.org/10.1111/jols.12372>
- Mufwene, S. S. (2020). Decolonial linguistics as paradigm shift. In A. Deumert, A. Storch, & N. Shephard (Eds.), *Colonial and decolonial linguistics: Knowledges and epistemes* (pp. 289–300). Oxford University Press.
- Murphy, K. R., & Aguinis, H. (2019). HARKing: How badly can cherry-picking and question trolling produce bias in published results? *Journal of Business and Psychology*, 34, 1–17. <https://doi.org/10.1007/s10869-017-9524-7>
- Nagle, C., Baese-Berk, M., Amengual, M., & Casillas, J. V. (2024). *Sound communities: A quantitative proposal for studying bilingualism in context*. PsyArXiv. <https://doi.org/10.31234/osf.io/m67tx>
- Nosek, B. A., Alter, G., Banks, G. C., Borsboom, D., Bowman, S. D., Breckler, S. J., et al.others. (2015). Promoting an open research culture. *Science*, 348(6242), 1422–1425. <https://doi.org/10.1126/science.aab237>
- Quintana, D. S. (2020). A synthetic dataset primer for the biobehavioural sciences to promote reproducibility and hypothesis generation. *eLife*, 9, 1–12. <https://doi.org/10.7554/eLife.53275>
- Roberts, S. O., Bareket-Shavit, C., Dollins, F. A., Goldie, P. D., & Mortenson, E. (2020). Racial inequality in psychological research: Trends of the past and recommendations for the future. *Perspectives on Psychological Science*, 15(6), 1295–1309. <https://doi.org/10.1177/1745691620927>
- Simmons, J. P., Nelson, L. D., & Simonsohn, U. (2011). False-positive psychology: Undisclosed flexibility in data collection and analysis allows presenting anything as significant. *Psychological Science*, 22(11), 1359–1366. <https://doi.org/10.1177/0956797611417632>
- Singh, L., Killen, M., & Smetana, J. G. (2023). Global science requires greater equity, diversity, and cultural precision. *APS Observer*, 36. Retrieved from <https://www.psychologicalscience.org/observer/gsequity-diversity-cultural-precision>
- Stieglitz, S., Wilms, K., Mirbabaie, M., Hofeditz, L., Brenger, B., L'opez, A., & Rehwald, S. (2020). When are researchers willing to share their data?—Impacts of values and uncertainty on open data in academia. *PLoS One*, 15(7), 1–20. <https://doi.org/10.1371/journal.pone.0234172>
- Tsikewa, A. (2021). Reimagining the current praxis of field linguistics training: Decolonial considerations. *Language*, 97(4), e293–e319. <https://doi.org/10.1353/lan.2021.0072>
- Wicherts, J. M., Veldkamp, C. L., Augusteijn, H. E., Bakker, M., Aaert, R. C. van, & Assen, M. A. van. (2016). Degrees of freedom in planning, running, analyzing, and reporting psychological studies: A checklist to avoid *p*-hacking. *Frontiers in Psychology*, 7, 1–12. <https://doi.org/10.3389/fpsyg.2016.01832>
- Zwaan, R. A., Etz, A., Lucas, R. E., & Donnellan, M. B. (2018). Making replication mainstream. *Behavioral and Brain Sciences*, 41(e120), 1–61. <https://doi.org/10.1017/S0140525X17001972>