

## DATA POLICY

At *Cognitive Linguistics*, we strongly believe that research data should be made widely available to the research community in order to demonstrate the robustness and validity of the research presented in our journal, to encourage replication of published results, and to provide the community with opportunities to learn. We believe that such transparency improves the quality of science and benefits not only the wider research community but the researchers as well by increasing their impact and enhancing their citation rates.

For all manuscripts submitted as of July 2020, *Cognitive Linguistics* requires, as a condition for publication, that all data (and related metadata) and any code supporting the results presented in the paper should be made publicly available, at the latest at the time of acceptance. Exemption may be granted by the Editor-in-Chief, for example, in case of sensitive data. Research data should be [FAIR](#) (Findable, Accessible, Interoperable, Reusable), should be deposited in an appropriate open repository and needs to be assigned a persistent identifier and an appropriate license specifying the conditions for reuse.

### Policy summary for authors

#### Required

- Sharing of research data via repositories
- Data availability statements

#### Optional

- Data citation
- Prepare and share Data Management Plans

The journal will ensure that its policy on data sharing is adhered to, including helping reviewers and readers access research data where needed. For some papers, the decision to publish will be affected by whether or not authors share their research data.

### Definition of research data covered in the policy

All data necessary to replicate your study need to be made available. This refers to data produced by the authors (“primary data”) as well as to data from other sources that are analysed by authors in their study (“secondary data”). It may include (but is not limited to) spreadsheets of annotated data, questionnaires or test answers and the computer code used to analyse it. Your data should be accompanied by rich metadata, i.e., who collected the data, how, why, when, where, who funded the data collection and so on (e.g., in a format of a READ-ME txt file, or a Data Management Plan). In addition to that, you may share any other data relevant to your study that you think other researchers may find useful. This includes images, audio, documents, video, maps, raw and/or processed data.

### Definition of exceptions

Research data that are not required to verify the results reported in articles are not covered by this policy. Exemptions may be granted by the Editor-in-Chief especially in cases where publishing the data would be in conflict with personal data protection or intellectual property rights, or could threaten the completion of a larger research project. Even if the data themselves cannot be shared publicly, researchers should publish at least descriptive metadata – either via a repository, or as Supplementary Material to the article – and should consider alternatives to public sharing (e.g., via a deposition in controlled access repositories, anonymization or de-identification of data before public sharing, or sharing upon reasonable requests).

## **Embargoes**

Embargoes on data sharing are only permitted with the agreement of the Editor-in-Chief. If your data cannot be made available upon publication (for example, if the dataset cannot be published before the completion of a funded research project), the Editor-in-Chief may allow for the dataset to be published with an embargo period - that way, the data would be safely deposited in a repository, but others will not have access to the data until the time specified in the embargo.

## **Supplementary materials**

Researchers are strongly discouraged from using Supplementary Materials to publish their data as that would restrict access to the datasets.

## **Data repositories**

Unless your institutional or funder's policy requires otherwise, you should make your data and code available via a subject specific data repository. Subject specific repositories are generally better equipped to meet the needs of the community and your data will easily reach other researchers in your field. We highly recommend you publish your dataset in either [LINDAT/CLARIAH-CZ](#) or [TROLLing](#) (The Tromsø Repository of Language and Linguistics): Both repositories publish linguistic data, provide persistent identifiers for your datasets, support licensing, provide citation guidelines, and allow for an embargo period if that is necessary. Another platform frequently used in linguistic research is [OSF.io](#) (Open Science Framework), which also allows anonymised data sharing for the purposes of peer review. Computer code or software may be shared via other appropriate routes such as GitHub.

## **Data citation**

The journal encourages authors to cite any publicly available research data in their reference list. References to datasets (data citations) must include a persistent identifier (such as a DOI). Citations of datasets, when they appear in the reference list, should include the minimum information recommended by DataCite and follow journal style.

## **Data licensing**

Your data should come with an open license of use, so that other researchers know what they can or cannot do with your data (for example, [Creative Commons 4.0](#) CC0 or CC-BY are widely used for such purposes). Your data should be assigned a persistent identifier (e.g., DOI) to ensure the link will be preserved even if the dataset moves to another website. The journal does not enforce particular licenses for research data, where research data are deposited in third party repositories. The publisher of the journal does not claim copyright of research data.

## **Data availability statements**

Your submitted manuscript must include a Data Availability Statement (DAS), which should be added as an unnumbered section called "Data availability statement" between the Acknowledgements and the References. In the DAS, you should specify (in the future tense) how the supporting data will be made

available, or explain why the data (or subset of data) cannot be made available and specify how an individual researcher may obtain access to the unpublished data, if at all possible.

Upon final acceptance, the Data Availability Statement must be updated to reflect how the supporting data can be accessed (via a permanent link), or why the data (or subset of data) cannot be made available and specify how an individual researcher may obtain access to the unpublished data, if that is possible.

Data availability statements can take one of the following forms (or a combination of more than one if required for multiple types of research data):

- The datasets generated during and/or analysed during the current study are available in the [NAME] repository, [PERSISTENT WEB LINK TO DATASETS]
- The datasets generated during and/or analysed during the current study are not publicly available due to [REASON WHY DATA ARE NOT PUBLIC] but are available from the corresponding author on reasonable request.
- Data sharing is not applicable to this article as no datasets were generated or analysed during the current study.
- All data generated or analysed during this study are included in this published article [and its supplementary information files].
- The data that support the findings of this study are available from [third party name] but restrictions apply to the availability of these data, which were used under license for the current study, and so are not publicly available. Data are, however, available from the authors upon reasonable request and with permission of [third party name].

## **Data formats and standards**

*Cognitive Linguistics* encourages authors to share research data using data formats and standards recognised by their research community. Please see [FAIRsharing.org](https://fairsharing.org) for more information on established data sharing formats and standards. The journal prefers research data to be shared in open file formats – those that do not require proprietary software to access - where possible. For example, tabular data should be shared as CSV files rather than XLS files.

## **Mandatory data sharing**

Your research data must be made available at the time of formal conditional acceptance at the latest. Your paper will not be published until it includes information on the accessibility of the data and any code used to analyse it. When our Editor-in-Chief decides to accept your paper for publication, you will receive a decision letter asking you to de-anonymise your manuscript, conform to the journal stylesheet and publish your data. Your final revision should thus include a permanent link to the dataset. You may, of course, publish your data earlier to allow reviewers access to your data.

Note that, for Registered Reports, data need to be made available upon submission of a Stage 2 manuscript. For more information on Registered Reports, check the journal website.

### **Research data and peer review**

Peer reviewers are encouraged to check the manuscript's Data Availability statement. Where applicable, they should consider if the authors have complied with the journal's policy on the availability of research data, and whether reasonable effort has been made to make the data that support the findings of the study available for replication or reuse by other researchers. The Consulting Editor for Methodology is entitled to request access to underlying data (and code) when needed for them to perform their evaluation of a manuscript. Peer reviewers may request the underlying data (and code), which the authors will need to provide, but will be asked to sign a confidentiality agreement.

### **Data Management Plans**

The journal encourages authors to prepare Data Management Plans before conducting their research and encourages authors to make those plans available to editors, reviewers and readers who wish to assess them.

### **Researcher/author support**

For any questions regarding the data policy, please contact our Editorial Office at [cogling@degruyter.com](mailto:cogling@degruyter.com)