COMMENTARY

Peer review in linguistics journals: Best practices and emerging standards

ALEXANDRA D’ARCY                               JOSEPH SALMONS

University of Victoria                University of Wisconsin – Madison

We present new data on peer review practices in linguistics journals, reporting the results of an online survey of editors. This paper aims to increase understanding of the processes and practices of peer review for everyone involved—editors, authors, reviewers, and readers. Apprehending concretely how peer review happens from beginning to end and how editors think about it should help to demystify the process, especially for graduate students and early career researchers, and make the experience somewhat less stressful across the board. Editors, authors, and reviewers all share, we trust, a desire for high professional standards and best practices. We hope to stimulate further discussion of these issues in the field and development of field-wide standards.*

Keywords: academic publishing, professionalization, reviewing, editing, editorial practice

1. Introduction. Like all active scholars, we know well from personal experience what it feels like to navigate the peer review process, from the choice of where to send a paper and preparing a manuscript to the wait for an editorial decision and then on to the full range of reactions according to the outcome: the pain and vexation of a rejection, the sometimes difficult engagement with changes called for in a ‘revise and resubmit’, and, yes, the occasional joy of an acceptance with modest required revisions. A remarkable degree of academic experience is shaped by peer review. Yet as long-time editors ourselves, we are intensely aware that there is an acute lack of reliable information about how peer review works in our field. We have too often seen frustration from colleagues, junior and senior, at opaque processes and outcomes that seem bewildering and are sometimes experienced as ad hoc, even unreasonable. The challenges come to bear in particularly crucial ways as faculty and other mentors work hard to bring along students in understanding how to navigate peer review. All of this led us to do what linguists do: gather and analyze data on a problem. We have engaged in years of discussions with other editors, reviewers, and authors, both experienced and soon-to-be. Ultimately, this led us to design and conduct a survey of editors of linguistics journals, the results of which we have discussed with many editors, colleagues, and students. That is, we have qualitative and quantitative data. Pieces of what we present here will align with intuitions, understandings, and experiences of many, but we gather from discussions of earlier drafts of this article that some pieces will come as surprises to everyone or almost everyone who reads these pages. We are creating baseline evidence not just to inform the field but also to improve the way we as editors, reviewers, and authors do our work, to illuminate sometimes dark corners of the process for students and early career researchers, and to motivate future work to move beyond this start.

As a professional association, the Linguistic Society of America (LSA) has long been concerned not only with the scientific study of language (i.e. academic and intellectual...

* We thank members of the Committee of Editors of Linguistics Journals for sharing their practices of peer review with us, in the survey and far beyond it. This paper has been shaped by discussions and comments from many people, and we are grateful to Lubie Alatriste, Claire Bowern, Ewa Czaykowska-Higgins, Robin Dodsworth, Katie Drager, Peter Grund, Mirva Johnson, Lex Konnelly, Jorge Rosés Labrada, Samantha Litty, Monica Macaulay, Onowa McIvor, Christine Mallinson, Carolina Amador Moreno, David Natvig, Panayiotis Pappas, and Tom Purnell for providing feedback and sharing their reflections with us. We also thank Sophia Chan, who provided keen research assistance, and Denae Dyck, who provided meticulous copyediting support. We benefited greatly from feedback from Andries Coetzee and John Beavers. The usual disclaimers apply.

pursuits) but also with professional components of such study. A number of committees, standing and ad hoc, are dedicated to those aspects of linguistics. The Committee of Editors of Linguistics Journals (CELxJ) is one such group, open to all editors of linguistics journals, broadly conceived. Members of CELxJ regularly engage in debate, discussion, and professionalization regarding academic publishing. While efforts are typically public-facing (e.g. workshops, symposia), conversations about concerns and practices tend to be more localized, occurring at the annual CELxJ meeting or on the group’s listserv. This article reports experiences of the committee’s membership—collected via an online survey—more broadly, with the aim of demystifying a core practice of journal publication: peer review. At the same time, we aim to understand and promote best practices and professional standards. Best practices can be defined as ‘procedures that are accepted or prescribed as being correct or most effective’ (OED).1 We understand such standards to include, for example, accountability to data, appropriate modeling (theoretical, statistical), and inclusive, accurate, and thorough citation practices, alongside critical and constructive engagement with authors’ perspectives and arguments.

Peer review, the assessment of scholarly work by subject-matter experts, is the standard in linguistics, as elsewhere in the academy. Indeed, in linguistics, whether for squib, chapter, article, or monograph, peer review undergirds research dissemination, designed to ensure the calibre and integrity of published work (through assessment of theoretical, empirical, and descriptive adequacy) and to improve quality (through suggested improvements to rhetoric and argumentation).2 It is also used to assess questions of fit and suitability for the platform.

The definition of ‘peer’ is coming under increasing and overdue scrutiny, as scholars begin to push back on the gatekeeping that takes place (e.g. McIvor & Rosborough 2020). At the same time, alternate models of how to engage in effective review are forthcoming (open review, transparent review, collaborative review, postpublication review, dynamic review; e.g. Ware 2008, Kelly et al. 2014, Björk & Hedlund 2015). Nonetheless, ‘peer review’ itself is likely to remain entrenched. Although overall satisfaction rates hover in the 60–70% range across disciplines (Ware 2016), peer review is broadly supported by academics at rates that are consistently over 80% of respondents (e.g. Bornmann 2011, Ware 2008, 2016).

As supervisors, authors, and editors, we share this position. Indeed, we have become staunch advocates for peer review, having seen the results firsthand, both in our own work and in the work we oversee through our various professional roles and responsibilities. In addition to being widely supported in academic circles as a quality assurance mechanism, peer review is valued by authors for the improvements they believe it brings to their published research (Ware 2008, 2016). At the same time, it is worth spelling out the assumptions and culture-boundedness of this stance as it relates to linguistics journals. Specifically, it is situated in particular conceptions of what academic work is for (advancement of science), who it is for (other linguists), and what it produces (academic research articles). This is neither the only model available nor the most

---

1 Oxford English dictionary online; consulted February 2021.

2 In some fields, especially those centered on community-engaged scholarship, peer review is an important presubmission step, functioning as a measure of accountability and integral to understandings of reciprocity: it is designed to ensure that community goals and wishes are respected and that cultural knowledge is handled in a good way and in an affirming way. We acknowledge the importance of this to the field specifically and to the production of decolonial, intersectional, transnational, and social justice-oriented knowledges generally (see e.g. Smith 1999, Kirkness & Barnhardt 2001, Rice 2006, Czaykowska-Higgins 2009). Research is not acultural.
appropriate one for all fields and kinds of scholarship (e.g. Mihesuah & Wilson 2004). We increasingly often do our work not only with other scientists but also with language communities and/or the public in mind. This diversification of who the work is for reflects refinements to ethical models on the one hand and the interdisciplinarity of our field and its impacts and implications outside the academy on the other. We need to reevaluate our models and expectations of peer review as a result.3

We also recognize that, despite overarching reliance on and support for peer review, it is the source of regular debate and disagreement among scholars, and it suffers from systemic challenges. Complaints about timelines, workloads, and the tenor of reviews are well attested (Ware 2008, 2016). There are concerns about its potential to stifle creativity, and its efficacy at catching errors is sometimes questioned (e.g. Schroter et al. 2004, Schrotet et al. 2008). There are also innate biases—social and systemic—in a process that functions as a gatekeeping mechanism (e.g. Link 1998, Budden et al. 2008, Cronin 2009, García et al. 2015, 2016, Helmer et al. 2017, Lanehart & Malik 2020), alongside general concerns of what Bal (2018) refers to as ‘turf policing’, such that innovation can be resisted when it pushes against established positions and knowledge systems. Finally, it can reproduce or entrench assumptions about what kind of work is valuable, and what kind is not (e.g. McIvor & Rosborough 2020). Our purpose here is not to debate these issues, each worthy of attention in its own right, but to review the state of the art of peer review in linguistics journals, a discussion that necessarily leads us into best practices and ways to improve our outcomes. In other words, we seek to illuminate processes and practices of peer review as they relate to linguistics and to uncover points where we can do better.

In §2, we sketch our methods and data, and we present results in §3, covering the process from submission to decision. We dedicate considerable attention to reviewer selection and review content, elements of peer review that shape outcomes. Section 4 notes effects of the COVID-19 pandemic on the review process. We offer a brief discussion and conclude in §5. We do not simply report survey results but integrate those with our own experiences and what we have learned from other editors, authors, and reviewers.

2. Method and data. In late 2018, after surveying the literature on peer review and reflecting on discussions on Linguistics Twitter about the trials and tribulations of academic publishing, we developed a Qualtrics survey and sent a link to the list maintained by CELxJ. Although the editor’s committee is part of the LSA, the list is not restricted to the society’s membership. Indeed, it reaches a large and diverse international group of editors. We solicited quantifiable data and open comments from this group (see the supplementary materials at http://muse.jhu.edu/resolve/141 for the survey). The questions were intended to probe current editorial practices for linguistics journals as well as to engage editor reflections on key elements of praxis. In this regard, the survey is somewhat distinct within the broader literature on peer review. Quantitative surveys on experiences and beliefs about peer review are not uncommon (e.g. Ware 2008, 2016), but surveys directly engaging editors within a field, seeking both quantitative and qualitative data, are rare; we are unaware of research that explores the issues in this way.

We received forty responses from editors, not only of journals but also of book series. Many of those editors work with more than one journal, often as lead editor for one and

3 As colleagues noted in correspondence, peer review is implicated in establishing and reinforcing what linguistics is and who it is for, alongside who linguists are (or should be). This is all the more reason to interrogate and reevaluate models and expectations of our practices.
associate for another. This response rate comprises approximately one quarter of the list and about half of the journals that were actively represented at the time of the survey. Importantly, this engagement represented a wide range of journals. 70% (n = 28) are specialist journals, including most major subfields and many language- or family-specific platforms, and 30% (n = 12) are generalist journals, many familiar to most linguists. Participating editors also represent a wide geographic reach, reflecting different academic traditions. Beyond the survey, we benefited from extensive correspondence with some respondents and other editors.

To maintain the anonymity of the editors who completed the survey (a condition of participation), we do not reveal the journals included in the results. When we do provide specifics in our discussion, we do so with explicit permission from the journals—typically ones with which we have editorial experience ourselves or for which the editor shared information publicly elsewhere and approved its inclusion here. Displayed in Figure 1, the editorial experience of those who engaged with the survey and in discussion with us is deep. The median length of experience exceeds nine years (n = 40; Q6). Nearly half of the respondents had more than ten years of experience (45%, n = 18), with another substantial cohort reporting four to six years in an editorial role (22.5%, n = 9).

Given the breadth, scope, reach, and experience of respondents and responses, we are confident that the information we report captures current peer review praxis in linguistics journals.

3. THE PEER REVIEW PROCESS. In this section, we walk through, from an editor’s perspective, the progression of a manuscript once submitted: the basic setup and context in which review takes place (§3.1), how reviewers are selected (§3.2), what editors understand as good and bad reviews (§3.3), the timeline of the review process (§3.4), and how decisions are made and what happens with them (§3.5). The final subsection (§3.6) addresses remaining issues that emerged from the editors’ qualitative comments. Throughout, we report the total number of responses (n) only when it is fewer than forty.

3.1. CONTEXT. Linguistics journals typically have one or more editors, often some associate editors, and, generally, an editorial advisory board. Some have editorial assistants, managing editors, or other roles (a glossary of terms appears in the appendix). Be-
yond the fact that an editor is in some sense in charge, there is vast variation and there are no established norms. For example, being on an editorial board may be purely honorary—a name listed in the journal and on its website—or may involve an obligation to review four or five manuscripts per year.

Within living memory, submission to a journal involved sending multiple paper copies of a manuscript through a postal service to an editorial office, where it was processed by hand. Today, the process is almost exclusively electronic. Most recently, online editorial systems like Editorial Manager, Manuscript Central, and ScholarOne have rapidly established themselves as the norm, with 72.5% of respondents using them \( (n = 29) \). As with many other transitions from analogue to digital, this one was initially a challenge for many, judging from our experiences and anecdotes we have heard, but respondents generally like their systems, whether moderately or greatly (56%, \( n = 19 \)); only one reports disliking theirs.\(^4\) The remaining journals rely on email for all stages of submission (\( n = 11 \)).

These systems often include questions that require authors to confirm expressly that the work is original and not under review elsewhere. Outright plagiarism and recycling of already-published material appear to be relatively rare. We cannot stress enough, however, that submitting a manuscript to more than one journal at once is a serious matter. Double submission represents a substantive breach of professional ethics; some editors are inclined to not consider future submissions from authors who double-submit. Aside from creating unnecessary, redundant work across journals, it badly exacerbates existing pressures on reviewers and on editors (see §3.4). Double submission is also surprisingly easily discovered. Different journals naturally often converge on decisions about the best reviewers for a given manuscript, and it is standard practice for reviewers to report that they have been asked to review the same manuscript by another journal. Editors and editorial team members are by definition professionally active scholars who are asked to review manuscripts for other journals, including ones that have been submitted to their own journals. It is, unfortunately, common to discover that a manuscript has been doubly submitted, including explicitly dishonest claims that the manuscript is not under consideration elsewhere. We have both experienced this within weeks of writing this manuscript, in fact. We hasten to add that communication about specific submissions among editors of different journals must by definition be exceptional and handled with delicacy; editors too are obligated to maintain confidentiality regarding manuscripts under normal circumstances.

Upon submission, preliminary checks relating to manuscript completeness, legibility of tables and figures, anonymization of author(s), and suitability for the journal are typically performed by an editorial assistant and/or (associate) editor(s). Most journals now have word limits on submissions, and length is often part of the check, though as electronic publishing increases and paper-based submissions decline, some journals are relaxing length restrictions for those cases where the analysis justifies a longer article. Unsuitable manuscripts are desk-rejected, that is, declined without peer review. This decision is made by an (associate) editor, and a common and important part of desk rejection is to suggest other possible venues. For example, a manuscript that speaks over-
whelmingly to specialists in French might be rejected from a generalist journal with advice to submit to a Romanist journal because it is framed too narrowly for a broader readership. Beyond issues of basic quality of method, argument, and evidence, some journals desk-reject manuscripts based on lack of references to current work (in some cases, even in the journal to which the manuscript was submitted) or lack of original research (i.e. a focus on literature review without sufficient new analysis, etc.). The most common reason to desk-reject is that the manuscript deals with a topic outside of the journal’s mandate. To avoid desk rejections, some journals invite queries from prospective authors to check whether a manuscript is likely to fit. Most journals provide mission statements on their webpages and in the front matter to print copies, but reading articles published in a particular journal often yields the best sense of what kind of material it publishes, as well as what kinds of evidence, methods, and presentations appear in its pages. It is also worth talking to people who have published in a particular journal. For authors, a desk rejection is often a reminder to consider carefully what venue is most appropriate for any given manuscript.

We turn now to manuscripts that do fit a journal’s scope and mission, those that are sent out for review.

3.2. Selecting reviewers. Selection of reviewers is perhaps the pivotal moment in review. Editors report soliciting two (62%) or three (35%) readers per manuscript, while one, for a specialist journal, approaches four or more (2%) (Q10). Some include an editorial team member or members of their editorial board on manuscript reviews (32%), though this is concentrated within specialist journals. It is more common to involve members of the editorial team or board only for complicated situations, like conflicting reviewer reports or when reports are undelivered (66%, n = 38) (Q35).

The standard is overwhelmingly toward double-anonymous review (reviewers and author(s) do not know each other’s identities), reported by the vast majority of respondents (87%). Single-anonymous review (identities of reviewers are unknown to author(s)) was reported by just a handful of journals (13%) (Q12). Within this latter group, which includes a book series, it is the norm to involve members of the editorial team in decisions. Notably, although it exists, no one reports using deanonymized review as part of their standard, prepublication review process (exceptions may exist for special issues, however). Likewise, no one reports triple-anonymous review, where assignment of manuscripts to editors is done by an independent party so that even the responsible editors do not know the identity of authors or reviewers. Some journals, such as Frontiers, use ‘transparent’ peer review, such that the names of reviewers who endorse publication appear on the published article (the names of reviewers suggesting that an article be rejected are not made public).5

Some journals invite authors to suggest potential reviewers. Whether editors act on these suggestions is constrained by such factors as whether or not the potential reviewer has an existing conflict of interest with the author(s), has an article of their own under review at the journal, has recently reviewed something else for the journal, or is seen by the editor as providing sufficient relevant expertise. Reviewers with reputations within the editorial team for unhelpful, unconstructive reports tend to be avoided. In short, while some of the factors that inform this decision would be known to the author(s), such as conflicts of interest, others are confidential and known only to the editor(s). It is also possible for author(s) to request that a certain reviewer not be invited. This is not an option

5 A good overview of models of peer review can be found in COPE Council 2017:4.
that is often invoked, but when it is, best practice is to include a succinct and professional explanation (e.g. known different theoretical stances on the phenomenon in question). In our experience, editors take such requests seriously and give them due consideration.

To our surprise, most journals do not provide reviewers with specific instructions on how to structure the review (62%), though the remainder (37%) do (Q33). Only a few have specific checks on data, methods, statistics, and other technical aspects of papers. Where this is part of the journal’s practice, it includes some checking of the data and having a professional statistician on the editorial team. One journal emphasizes open data citation.

It is clear from the responses that standard practice is to rely on experts for quality assurance: almost all editors choose reviewers with an eye to controlling these matters (Q24; n = 37). We asked whether editors ‘expressly elicit a range of perspectives from reviewers of a given paper’, to make sure that empirical, methodological, or theoretical angles, for instance, are covered (Q21; n = 38); see also above on quality assurance. Most editors do this, and others seem to keep an eye on having distinct outlooks on a particular manuscript. Identifying such reviewers rests heavily on various kinds of institutional memory, such as long-serving editors (see Fig. 1, §2) or communication within an editorial team. This has the potential to hamper innovation, but editors can take active steps to curtail such effects by drawing in reviewers at a range of career stages, which diversifies reviewing lenses and brings along new colleagues with new perspectives.

It is in these respects—selection of and reliance on reviewers—that deep editorial experience is particularly valuable. The longer one works in some editorial capacity, the more familiar one becomes with both individual areas of research strength and individual reviewer practices. This view of the field enables editors to solicit from an increasingly broad and specialized pool of experts, and facilitates reliance on reviewers for checks on empirical, theoretical, methodological, and descriptive aspects of scholarship. We see this richness of editorial experience and expertise as a strength, but as noted (§2, §3.1), it is not unproblematic. Efforts to develop more systematic checks, beyond reviewer expertise, may be an area of change in the near future.

3.3. Good and bad reviews. Reviews are by nature advisory. For authors and editors alike, some reviews are better than others. By ‘better’ we mean more constructive and helpful, guiding and supporting both editorial decisions and authorial revisions. That is, some reviews do more to shape better published research. There is substantial variability in the length of reviews, with roughly half of responding editors indicating that they are typically one to two pages in length (53%, n = 38), and the remainder reporting that reviewer reports are typically longer than two pages (47%) (Q20). Reviews can be less than half a page or may exceed eight pages, reflecting a diversity of expectations, understandings, and levels of engagement when linguists engage in peer review.

Beyond the reports themselves, reviewers also provide confidential feedback to editors by email or through built-in features of reviewing software. We asked editors about their experiences with this, trying to assess the ways in which this practice is used to support the review process (Q31). For some, comments to editors are rare, but others indicate that this option is invoked often. The most common reason is to give blunter assessments than the reviewer wants to convey in their report, so that the editor sees clearly the level of concern reviewers have about a manuscript. Others use it to alert editors to potential malpractice, such as plagiarism, to disclose potential conflicts of interest, or to provide relevant history, such as having reviewed an earlier version of a paper for another journal. Where reviewers are asked to give an explicit recommendation on the editorial decision, feedback provides reviewers an opportunity to tell editors that
they were on the borderline between two categories, such as revise and resubmit versus reject, and to outline their reasoning for having made the recommendation they did.

There are many aspects of academia in which we receive little formal training, including peer review (but see Sachs 2020). Around the time our survey was circulated to editors, we solicited feedback from linguists on Twitter about whether they were ever explicitly taught how to write a review (Figure 2). The resounding response was ‘no’. Most linguists learn how to review on the fly, using reviews they have received and advice from supervisors and other mentors as guideposts along the way. Some graduate seminars include mock review assignments, and some supervisors co-review manuscripts with students to support their professional development (see §3.6). Others co-author with students, providing models for writing, working through review content, and responding to reviews. Nonetheless, the overarching response was that reviewing is self-taught. It is no surprise then that there is great variability in the calibre of peer review. This is an area where professional societies, departments, and programs can develop professionalization strategies.

We asked editors what they consider the key elements of a good review to be (Q29), alongside how they believe reviewers can most effectively help both authors and editors (Q30). There is overlap in these questions and the answers, though the former probes for criterial components and the latter for supportive components. Across the sixty-three responses to these two questions, clear themes emerge. We discuss those first, before moving to unhelpful review content (Q34, \(n = 16\)) and what editors see as reviewing mistakes (Q32, \(n = 34\)).

Editors value reviews that make an attempt to understand an author’s goal and perspective (even if it differs from their own) and provide specific, helpful feedback. As summarized by one respondent, ‘a good review is written both for the author and for the editor’. Such reviews summarize the main argument, assess both strengths and weaknesses, recognize the potential of an article and guide the author(s) to develop it, and make clear suggestions for improvement (e.g. as noted by one editor, they outline how the evidence can be ‘better mobilized to support the argument’). Editors report that effective reviews ‘concentrate on how to improve the manuscript, not on why it is bad’, and they offer ‘not just criticisms, but suggestions for overcoming them’. They also provide full references for missing literature or for work that will help the author(s) moving forward. In short, helpful reviews are constructive and fair. They praise what works, and they provide supportive and concrete advice for strengthening the paper.

One editor stresses that all of this is important regardless of the suggested outcome, but it is especially so if the manuscript is likely to get rejected. Constructive guidance
will help the author revise for potential submission elsewhere. This is true for authors at all career stages, but particularly for work by students or early career researchers. A reject outcome is never easy to give or receive. It can, however, have a positive impact on research if the reviews and the decision letter outline productive and supportive directions for moving forward. As an author of a rejected manuscript once wrote to us (Anonymous, p.c.), ‘We all know rejection is usually part of the job, but there’s no reason it can’t be positive too’. One editor observed that ‘a review from an experienced researcher gives the author access to expertise, and a helpful review is generous in this respect’. We agree that generosity and balance are key.

Unhelpful components of reviews fall into three broad categories: overinvestment in minutiae, unsubstantiated assessments, and personalization of reviews. The first, minutiae, neither undermine nor devalue the review; they simply clutter it, making it harder for editors and authors to isolate substantive points of content. Specifically, editors find it unhelpful and unnecessary when more than passing attention is paid to matters of copyediting and style. One editor ‘would like to see reviewers focus on substantial aspects (content, arguments, theory, etc.) rather than commas, spellings, etc. The latter can be fixed and we don’t need to see long lists of “errors”.’ Respondents also noted that reviewers can become ‘bogged down’ in the details rather than focusing on the arguments themselves.6 A simple and common best practice is dividing a review into major versus minor issues, so that the serious matters of substance and structure are highlighted, while smaller ones are available for authors to draw on while revising.

The second unhelpful review component is when reviewers make sweeping statements (as vague as ‘there are a number of problems’) but do not provide specifics. Included here is a failure not only to suggest possible solutions when aspects of a paper can be strengthened or improved but also to outline strengths of a paper when the reviewer is positive about the manuscript. In other words, stamps of approval without context are as problematic as stamps of disapproval without context. As one editor notes, this is a source of difficulty in two respects—comprehending the paper and making a decision:

This isn’t helpful to the editor, who might not understand why the reviewer likes the paper or sees value in it, and it is probably one of the most common sources of splits between ratings. If one reviewer says ‘reject’ and the other says ‘accept’ but doesn’t provide any justification, it is hard to know what to do.

Crucially, this lack of guiding advice is as hard on authors as it is on editors, because it can create confusion or even misdirection in terms of moving forward. At the other end of the scale, an editor writes, ‘self-indulgent monologues of the referees’ own achievements’ that fail to achieve ‘a true engagement with the claims of the paper’ are also unhelpful. A recurrent theme is to provide constructive feedback, in clear language, that indicates to editors and authors what is successful, where refinement is needed, and how the author(s) could consider going about that.

The third category of unhelpful review content includes activities often ascribed to the mythical Reviewer 2, a sometimes contrary and at all times unhelpful reviewer: attempts to rewrite the paper in the reviewer’s own voice or style, to build the narrative or analysis they themselves would pursue, to include (extensive) citations of their own

6 A benefit of this is that it steers reviewers away from becoming entrenched (intentionally or otherwise) in language ideologies that are tied to the construct of ‘the standard’. We include here assumptions about language ability and raciolinguistic constructions of author identity. If parts of a paper are unclear, it is important to ask for clarification, but comments regarding ‘native speaker’ oversight are best avoided. Most issues can be resolved during copyediting.
work (i.e. beyond those directly relevant), and to soliloquize on their own point of view. As multiple editors note, reviewers should not attempt to make the paper ‘one that they would have written’. For all concerned, reviewers should honor the author’s perspective and should not expect the paper to be about something other than what the author intended—an approach that not only makes the review irrelevant but that also demoralizes junior scholars (because it dismisses their perspective(s)) and may even hamper innovation. And in a second review, after resubmission, reviewers should not raise issues they did not raise in the first round but could have—the only exception is when they discover a problem of substance that has emerged through revision. The reason is that it ‘moves the goalposts’ for the author.

Other practices in this third category are things such as personal attacks and inappropriately critical reviews. (Harsh and ad personam reviews may have become less common over time, but we still have a long way to go—they do still occur; see §3.5). In short, reviews should not be personal, aggressive, or rude. This seems self-evident, and yet this point was made at some point by nearly all of the respondents. These problems exist along a scale, some harsher than others. We have been told our work is ‘pretentious’, ‘bad’, ‘stylistically bizarre’, and ‘a letdown’, that it ‘devotes more thought to imaginative word selection than to conducting and reporting empirical research’. These examples illustrate reviewing behaviors that are better avoided, not only because they are unkind—there are gentler ways to make the same points—but also because they are unproductive. We discuss how editors handle these reviews in §3.5.

In sum, good reviews note not only problematic or underdeveloped aspects but also positive aspects of the paper. They offer substantive feedback, regardless of recommendation. They engage with the arguments as presented by the author, and they do so in the spirit of constructive criticism and collegial support. They approach the exercise as ‘an opportunity for development, not for gatekeeping’ (Catherine Anderson, @candersHamilton, April 5, 2021, Twitter). Bad reviews offer verdicts but do not substantiate them, or they become bogged down in details the reviewer wishes to see but do not provide a big-picture assessment. They attack, criticize in ways that are not constructive, and are unkind. They are also ungenerous, both in holding back substantive engagement and in reading incompetence into omission, assuming that a missed reference, for example, entails that the author is inept or inexpert, rather than giving the benefit of the doubt. In short, bad reviews fail to follow the mantra of reviewing ‘unto others as you would have them review unto you’ (Figure 3; @AcademicsSay, January 22, 2020).7

7 We endorse the advice from Brown (2015) on how to not be Reviewer 2: be the right person for the article, be in the right frame of mind to review, review as you want to be reviewed (see Fig. 3), and be specific and helpful.
In the end, authors who submit to peer-reviewed journals are going to get bad reviews. We hope that understanding better how editors think about these reviews helps authors process them.

3.4. The chronology of review. A pressing concern among all involved is the timeline of review: graduate students and early career scholars are under often intense pressure to produce journal publications, while reviewers are taking on unpaid and generally unrecognized labor that must be shoehorned into increasingly overstuffed work schedules. We devoted a number of questions to this.8

Journals that use an online management system typically give reviewers a week to respond to initial invitations, after which (automated) weekly reminders are sent (Q11, n = 28). For journals that do not use an online system, there is greater variation, with editors giving reviewers anywhere from three days to two weeks to respond, after which they either send one or two reminders or simply invite someone new to review the paper in question (n = 10). The mean rate of failure to respond to an invitation to review is 23% overall (Q16, n = 33). This figure is nontrivial, and the range is extraordinarily large, differing somewhat across specialist and generalist journals. For specialist journals, reviewers fail to respond at rates that run from 2% at one extreme to 71% at the other. This high end of the range is an outlier, one associated with a unique situation. If we set aside that particular platform (for which we had engagement with two editors, reporting nonresponse rates of 61% and 71%, respectively), the mean rate of failure to respond to an invitation is 21% (n = 22) for specialist journals (range 2% to 50%). For generalist linguistics journals, the percentage of invitees who are nonresponsive runs from 3% to 40%, with a mean of 20% (n = 9). Setting aside delays due to spam filters, failure to respond to an invitation to review is in essence gross negligence toward the author(s): it often delays decisions by weeks or more and can thus damage the careers of people under intense pressure to publish.

Typical timelines from first invitation to a confirmed set of reviewers range from a few days to a month, though we have seen it take substantially longer. Once an invitation is accepted, most journals give reviewers four to eight weeks to submit the review (69%, n = 39) (Q13). Just a handful have timelines that exceed ten weeks (10%), while an expected turnaround of two weeks or less is exceedingly rare (2.5%).

Editors report that only about half of reviewers submit their reports on time: the median is 55% (Q14, n = 35).9 The range is again extremely large, with estimates of timely submission varying from a low of 9% to a high of 90%. This does not correlate with how long reviewers are given—say, four versus eight weeks. Rather, timely submission rests on willingness or ability to adhere to the deadline. Of course, life happens. If for some reason a timeline that initially seemed manageable becomes a challenge, editors are generally willing to negotiate (reasonable) extensions. As a matter of professional courtesy, to editors and authors alike, we urge reaching out to editors when a deadline cannot be met.

At the same time, it is relatively rare for linguists to fail to submit a review once they have agreed to do one (Q15, n = 33). Thus, where timeliness of review submission is a challenge overall, ultimately meeting obligations in the end is much less of one. There

8 The responses to these questions were collected in prepandemic times; we discuss how the pandemic has affected multiple aspects of peer review in §4.

9 Reviews are less often submitted early. The median estimate was that just 24% of reviews that came in 'on time' arrived at least a week before the editor-determined deadline (Q17, n = 34). The remainder of 'on time' reviews are typically received within a day or two (either side) of the requested date.
is a split, however. The majority of editors report that less than 10% of committed re-
views fail to come in, with the mean nonsubmission rate sitting at 5% in this cohort
\(n = 23\). This is the low end, and it represents a substantial proportion of our responses
(70%). The remainder of editors, a small subset, report nonsubmission rates that range
from 15% to 50% \(n = 10\). Although this reflects a minority of peer review (30%), what
is troubling here is not only the high rates of submission failure on this side of the split
(mean 30%) but also that all but one of the journals in this category are specialist jour-
nals, precisely the venues where expert input is harder to get and more critically needed.

All of these factors combine with editorial workloads to yield a typical timeline, from
initial submission to editorial decision, of two to five months, though responses gener-
ally skew toward the lower end of this scale (81%, \(n = 37\)) (Q37). It is rare, desk rejec-
tions aside, to share a decision with authors within two months of submission (3%), and
it is atypical for a decision to take longer than five months (16%).

3.5. Decisions. When all reviews are in, most editors read a submission and so have
their own views on the strengths and points of necessary development of any individual
manuscript. Indeed, we take the stance that an editorial decision cannot be made re-
sponsibly if the deciding editor has not read the submission in detail. Still, the reports—
taken together—are the primary determinant of outcome. When assessments are in
agreement, the pathway to a decision is naturally clearer and easier. The literature sug-
gests that cross-reviewer agreement (i.e. interreviewer reliability) is low across the
humanities, social sciences, and sciences (e.g. Cicchetti 1991, Bornmann et al. 2010,
Bornmann 2011, Chibnik 2016). We asked three connected questions relating to agree-
ment across recommendations (Q27, \(n = 39\)), parallelism in qualitative components of
reviews (Q38, \(n = 37\)), and the expected degree of agreement and overlap (Q39, \(n = 15\)).
Tables 1 and 2 present the results for the first two questions in this series; responses to
the third are woven into our discussion of these results.

<table>
<thead>
<tr>
<th>RESPONSE</th>
<th>(%)</th>
<th>(N)</th>
</tr>
</thead>
<tbody>
<tr>
<td>More often than not, they make the same recommendation or the recommendations are similar (e.g. major revision vs. minor revision).</td>
<td>25.6</td>
<td>10</td>
</tr>
<tr>
<td>Their recommendations are generally quite similar.</td>
<td>59.0</td>
<td>23</td>
</tr>
<tr>
<td>It’s about 50/50 agree vs. do not agree (e.g. reject vs. minor revision).</td>
<td>14.4</td>
<td>6</td>
</tr>
<tr>
<td>There is relatively little agreement between reviewers.</td>
<td>—</td>
<td>0</td>
</tr>
<tr>
<td>More often than not, there is no agreement between reviewers.</td>
<td>—</td>
<td>0</td>
</tr>
</tbody>
</table>

TOTAL \(n\) 39

Table 1. Do reviewers generally agree on their recommendation (e.g. major revision)?

<table>
<thead>
<tr>
<th>RESPONSE</th>
<th>(%)</th>
<th>(N)</th>
</tr>
</thead>
<tbody>
<tr>
<td>An overwhelming amount</td>
<td>—</td>
<td>0</td>
</tr>
<tr>
<td>A lot</td>
<td>37.8</td>
<td>14</td>
</tr>
<tr>
<td>A moderate amount</td>
<td>59.5</td>
<td>22</td>
</tr>
<tr>
<td>Only a little</td>
<td>2.7</td>
<td>1</td>
</tr>
<tr>
<td>None at all</td>
<td>—</td>
<td>0</td>
</tr>
</tbody>
</table>

TOTAL \(n\) 37

Table 2. How much overlap, consistency, or thematic parallelism is there between reviews?

Although the questions in Tables 1 and 2 present a relatively blunt instrument for as-
sessing interreviewer reliability, responses indicate a broad tendency toward basic
agreement, more than we would expect from chance. Editors and authors are concerned
about inherent arbitrariness in the process, especially with regard to reviewer selection,
but these responses suggest that the problem may be smaller than some fear. Still, inter-
reviewer agreement is ultimately not the best proxy for whether or not the system works. While editors necessarily make decisions on a manuscript-by-manuscript basis, they do so in the context of all submissions, editorial policy, and broader goals of the journal and the field.

All editors deal with diametrically opposed reviews, for example, one urging acceptance and another rejection, but as one editor notes in response to the third question in this series, ‘blatant disagreements are under 5%’. This exceptionality is captured well by the survey responses, both in terms of recommendations and in terms of qualitative content. Another editor writes that they would ‘expect less agreement in general’ but that ‘80% of articles have very similar or complementary reviews’. Notably, the majority of studies of reviewer agreement in the literature focus on the recommendations (e.g. Cicchetti 1991, Bornmann et al. 2010, Bornmann 2011, Chibnik 2016). For large-scale quantitative analyses, this is natural and understandable. But there can be incongruities between an overall recommendation and the content within a review. And reviews may make different overall recommendations but raise very similar concerns. As one editor observes,

I find there to be more disagreement in the specific recommendation of a review (e.g. ‘revise’, ‘accept with revisions’, etc.) than in the implied recommendation that one gathers from reading the reviews themselves. For instance, a lot of reviewers will make a ‘minor revision’ decision when their review suggests ‘major revision’.

We understand this as signaling that it is the degree of alignment between the qualitative assessments in reviewer reports that is primary, not the degree of agreement between recommendations—a distinction that we ourselves regularly draw upon as editors.

Given the tendency of editors to select reviewers with distinct, complementary expertise (§3.2), we understand the trends in Tables 1 and 2 as positive, and they certainly ease both the editor’s job and the author’s job. As already noted (§3.3), reviews are by nature advisory, but it appears that editors are getting, in the main, relatively consistent or overlapping advice from different perspectives. Thus, although some scholars consider variance in reviewer agreement a factor that creates a ‘luck of the draw’ scenario (Esarey 2017:964), the picture that emerges here is one that instills some confidence in the outcomes that inform editorial decisions in our field.

The number of hands involved in the decisions themselves is subject to variability. Some editors handle editorial decisions alone (38.5%, n = 39) (Q9), while the same number report involving one other member of the editorial team. The remainder (23%) range from involvement of three people (n = 2) to more than five (n = 4). The trade-off is clear: involvement of more people means more checks for clarity, fairness, and balance, but it also takes more time and requires more work from more people. Broad involvement in decisions generally strives to reach consensus on outcomes, an effort that is facilitated when reviewers agree—which, as we have seen, is reported to be relatively common in linguistics.

The editor’s job at this point is to synthesize the reports in a meaningful way that makes it transparent for the author why the decision is what it is (e.g. a revise and resubmit) and what aspects of the manuscript the author should particularly attend to moving forward (e.g. reframe the argument, reconsider a part of the methodology, refine the interpretation). These skills, too, are acquired informally, though active editorial teams offer an opportunity to bring along new editors in the culture of a journal and subfield. This information should be explicit in the decision letter, providing a road map to help authors work their way through the reviews. In our experience, another reason editors provide context for decisions is that there are points in reports with which they do not
agree, which contradict each other across reviews, or which require some other kind of attention. It could be that a review suggests an approach the editor deems unnecessary, or that a review seems overly critical, yet its length and attention to detail signal reader interest and investment in the outcome, or that the tone of a review is unfortunate, and so forth. Decision letters can be longer or shorter, but they need to include substantive feedback and not simply decree a decision.

As discussed in §3.3, reviews may contain undesirable elements. Apart from their impact on authors, such reports pose problems in formulating constructive editorial decisions. We asked an open-ended question to assess how editors manage ‘bad apple’ reviews and the level of negativity they tolerate in reports (Q22). This elicited thirty-four comments, of which fifteen made some reference to how rare this is. Many comments were careful to distinguish between negative, critical reviews, an obvious part of a sound process, and reviews deemed ‘harmful’, containing ‘abuse’ or ‘excessive vitriol’. For problem cases, editors employ various strategies, beginning with avoiding inviting reviews from readers known to be problematic. It appears to be common practice to ban these reviewers once they reveal themselves, not only from reviewing original submissions but also from reviewing a resubmission they reported on in the first round. As one editor put it, ‘such reviewers self-select themselves out of the pool’. Some editors ask reviewers to tone down or rework problematic parts of reports, while others edit, excerpt from, or simply delete problematic reports, though others are uncomfortable with altering reports from outside readers. Another strategy is to solicit an additional, ‘Solomonic’, review in such cases, a practice that was explicitly noted by a handful of the respondents. We can also infer that it is fairly standard practice to use decision letters to comment explicitly on problematic aspects of reports, along the lines of ‘we disagree with Reviewer 1’s assertion that this additional statistical test is needed’.

To a question about whether editors have ever held a review back from an author, only a few had, with two noting that it had happened only once and a third writing ‘maybe once or twice in 20 years of editing and hundreds of papers’. This may be because particularly bad reviews are infrequent (as noted, most bad behavior occurs along a continuum, interwoven with helpful content), and/or because editors provide light editing before sharing them with authors, and/or because such reports are contextualized in decision letters. Some editors may simply send out all reviews, as received, regardless of tone, tenor, or content. In our view, even if rare, this latter practice should be halted. It serves no one, and may well do harm.

Increasingly common is for editors to share reviews and editorial decisions with reviewers, in part as a way of thanking reviewers for their work but also in the service of transparency. A potential added benefit is that reviewers have the opportunity to see what other reviews look like, providing an opportunity to learn something not just about the evidence and arguments at hand but also about reviewing—in terms of both practices to emulate and those to avoid. In our responses, 57.5% ($n = 23$) of editors share decision letters with reviewers (Q18), and 60% ($n = 24$) share all reviews with other reviewers (Q19). Perhaps expectedly, the correspondence between these two practices is very high: twenty of the represented journals share both decision letters and reviews with reviewers. Notably, sharing review materials and outcomes appears to currently be more common among specialist journals than among generalist ones (57% vs. 33%).

What does an author do when they get an editorial decision? Unless it is a particularly positive one, many people need a few days to digest and interpret the decision. If it is a rejection, it means looking for how to rectify the major problems and then find a more suitable outlet for the next iteration. Long gone are the days when it seemed acceptable
to immediately send a rejected manuscript to a new journal without serious revision. That
misses the fundamental benefit of peer review—to work through those aspects of the de-
cision and the reports that will help produce a better final manuscript. Rejection or ac-
ceptance, it is rare in our experience for the process not to yield valuable feedback.

Indeed, on initial submissions, an editorial decision is generally the beginning of a
new process rather than an end: the most common editorial decision is ‘revise and re-
submit’ of some form and often the assumed default decision on a good submission.
(And, yes, this goes for senior scholars too.) The survey data support this. On average,
just 33% of initial submissions are accepted (Q47, n = 33), but the range runs from 0%
to 76%. Journals with estimated initial acceptance rates at or above 50% are excep-
tional (n = 11). They are mostly specialist platforms, though distributionally, there is re-
ally no apparent difference across the kinds of journals that have these high acceptance
rates: 28% of specialist journals and 25% of generalist journals fall in this end of the
range. (Some journals also solicit contributions, which presumably aligns with higher
acceptance rates.) These high-acceptance journals are typically, but not categorically,
journals for which the editor (or a very small team) makes decisions. Most journals
have initial acceptance rates below 30% (58%, n = 19). As a result, when editors submit
their own work, most have come not to expect an acceptance upon initial submission
but hope to have useful input for improving the work to get an eventual acceptance. For
this reason, a common piece of advice to students and young colleagues is that an
‘R&R’ (as it is called) is a positive outcome. It signals that there is work to do but also
that there is a concrete path to publication.

This means working on revisions and writing a response to the decision. It is now
usual to request a response to the editor and reviewers, explaining how major issues
have been addressed, point by point. Even where a response is not expressly requested,
many authors provide one, and responses can give editors valuable information on the
revision. On occasion, it is appropriate or even necessary to push back on one or an-
other point, if a reader has been incorrect about facts, for example. Far more common is
that issues can be resolved by clarifying or explicating something in the text. Often, an
appropriate response to a point is simply to identify where the issue is dealt with in the
new version of the manuscript, such as ‘I’ve reworked §5.3 along the lines suggested’. Here,
too, a positive tone and gratitude are welcome, such as starting a response by
thanking the editors and reviewers for feedback.

As just noted, decisions and reports are increasingly shared with reviewers. To pur-
sue that further, we asked editors if, in the case that they request a letter from authors
about revisions they have made to address reviewer comments, that letter is shared with
reviewers who read the new submission (Q43). Of the thirty-four responses to this
question, a full 76% indicate they do (n = 26). In short, the reviewers of the revision
will likely see the author’s response to the decision and take it into account in evaluat-
ing the revised draft.

In our experience, there is a split in how authors respond to editor and reviewer feed-
back. Some mechanically address specific points from the feedback in what strikes us
as a principle-of-minimal-effort mindset. These are unsatisfying revisions, almost by
definition, and can be rejected. By contrast, editors relish revisions that embrace the

10 It may be helpful to remember that editors solicit reviews from relevant experts. If the manuscript was
not convincing to the experts who read the submission for one journal, it will likely not be convincing to those
who read the submission for another journal, as is.

11 We did not specify the inclusion or exclusion of desk rejections and do not know the rates of desk rejec-
tion at different journals.
spirit and not just the letter of the decision and reports. When an editor or reviewer suggests considering another perspective, with relevant references, the former approach may involve adding a couple of references and a comment to the effect that 'some people see things this way'; the latter involves reading the references (and often follow-on references) and considering how authors can sharpen the argument or better contextualize it. While this manuscript was under review, one of us worked with a graduate student doing their first peer review (with the editor’s permission) and the process yielded a revise and resubmit. In reviewing the revision, the student was visibly excited to see that the author(s) had embraced the full spirit of the report, saying basically, ‘I knew this could be a solid paper, but it turned out way stronger than I expected’. This young scholar now knows from personal experience the impact of a constructive review and the value of conscientious engagement with revision.

We did not ask about it, but some journals normally allow only one round of revision, typically making an up-or-down decision on the second iteration, while others may go through multiple rounds of revision and review. (Journals, including the LSA’s flagship Language, increasingly show the review history at the end of published articles, indicating how many times a paper has been revised.)

After careful revision, we expect acceptance rates to increase, and that seems to be the case. Editors estimate that at this stage in the process, accept outcomes (often with minor revision) rise from the 33% on initial submission to 73% on revision (Q47, n = 30). Here the range is from 18% (an outlier) to 95%, with the vast majority of editors reporting that acceptance rates are above 50% (80%). We interpret this as strong support for our assertion (§1) that peer review improves outcomes, a position corroborated by the literature more generally (e.g. Ware 2008, 2016). There is a moderate positive correlation between journals with high rates of acceptance at both stages and those with lower rates at both stages, which is to be expected ($r(31) = 0.51, p = 0.003$). More crucially, only two editors report that rates of acceptance do not improve markedly upon resubmission, both of which are for journals with exceedingly low levels of acceptance at both stages (generally in the range of 10%). We infer from this that earnest engagement with a revise-and-resubmit decision usually leads to eventual acceptance.

3.6. FURTHER ISSUES. When we asked editors what reviewers could do better (Q25), or what could be improved in the review process (Q26) (n = 55), a number raised issues related to professional courtesy. There was widespread acknowledgment that peer review relies on the generosity and goodwill of our colleagues, and that the system neither recognizes nor valorizes that generosity. In earlier stages of our careers, we both tracked all of our peer review activities in detail, including them in annual reports for salary assessment alongside applications for reappointment, tenure, and promotion. We strongly encourage early- and mid-career scholars to do the same, even if it is not required. Invitations to engage in peer review reflect a positive, if sometimes nascent, reputation. Normalization of such practices may facilitate an institutional sea change in higher education, such that peer review becomes more systematically recognized for the important scientific service it is.12

Returning to a point we just made, peer review relies on generosity and goodwill. Indeed, the most common reported reasons for engaging in peer review are intrinsic: to contribute meaningfully to one’s academic community, to help others improve their

12 We have not used Publons (a commercial online site for academic publishing activities) or Kudos (another, which tracks ‘impact’), but some publishers allow reviewers to opt in, enabling formal tracking of peer review and other activities for reporting purposes.
work, and to reciprocate the efforts made by others in review of one’s own work (Ware 2016). Editors noted that it is problematic and time-consuming when reviewers do not respond to invitations to review or when they fail to submit a review on time. A minimal level of professionalism (and collegiality) surely requires at least promptly letting an editor know that a review cannot be taken on. Editors also noted the challenges raised by colleagues who routinely decline invitations, as well as the practice of declining a request for a second review after having done the first review and recommended an R&R. This creates a kind of ‘double jeopardy’ problem, changing the evaluation set after revision. Editors also wished to see further engagement with the request to suggest alternate reviewers when declining invitations. Given that reviewers are selected on the basis of areas of research strength and expertise, this is purely pragmatic. Experts are best placed to know the others with similar areas of knowledge, and so making recommendations supports peer review practice in fundamental ways. It can also identify reviewers of whom editors were not aware—an important broadening of participating voices.

Reviewing timelines across academia are notoriously problematic, and linguistics is no exception (§3.4). One editor wishes that ‘reviewers would understand that if they want their own papers to be reviewed in a timely fashion, all of us need to be more generous in reviewing papers’. We are aware of at least two suggested best practices. One stipulates that scholars review as many manuscripts for journals as they submit in a given year. A second, related formulation became a mantra among the editorial team of *Diachronica* some years ago, thanks to Sheila Embleton. She suggested that scholars should review as many manuscripts per year as their manuscripts generate. So, if someone submits three manuscripts to journals in a normal year and the norm in the relevant area is three reviewers, they should review, on average, nine papers per year. We advocate this latter standard, as a matter of equity and fairness in a system that demands much of us, and we extend it to revisions for all concerned.13 The burden of reviewing normally becomes heavy only as scholars become established; in most subfields, graduate students and early career researchers are only occasionally asked to review, and the occasional review presents, in some sense, a learning opportunity, especially if editors allow or even encourage students to consult with their advisors along the way.

Two further concerns raised by editors pertain to anonymization and representation. Some editors worry that double-anonymous review leaves reviewers free to behave badly without repercussions (cf. §3.3, §3.5), while others note that it is becoming increasingly difficult to maintain anonymity and confidentiality. It is of course one thing to guess at an author’s or reviewer’s identity; it is quite another to undertake internet searches with the intent of making or confirming identification. If a journal uses a double-anonymous process, this latter behavior is unprofessional at best and unethical at worst—it is a violation of author anonymity (we return to ethics below). As already noted, new modes of review are emerging that arguably increase transparency and accountability. Preregistration of experimental research holds promise. Already the norm in some fields, it has been slower to catch on in linguistics, but a foothold has been established, especially in the form of registered reports (e.g. in research in first and second language acquisition, cognitive linguistics, and speech sciences; see, for example, Roettger 2021). We see this as an auspicious development.

---

13 Sheila Embleton describes the history of her formulation in correspondence: ‘Some of us were sitting around one day lamenting how hard it was to get faculty members to sit on tenure committees, which occur at several levels of the institution, when somebody asked how many committees one “should” sit on in order to in effect pay back for getting one’s own tenure. So I did the math for that one.’
Representation is a trickier, more layered issue. One editor noted that they were concerned about ‘getting into the system voices not usually heard’. This was echoed in various ways. For example, we heard that ‘there is a general idea out there that because the journal publishes in English, reviews have to be in English too, which limits participation’. We did not ask editors about practices of equity, diversity, and inclusion, though editorial teams are increasingly aware of systemic bias, not only in reviewer selection but also in qualitative aspects of reports. Some subfields remain skewed toward particular categories (typically cisgender white men), highlighting challenges within linguistics broadly. Other subfields are increasingly diverse, though we are unsure of the extent to which reviewing practices are impacted. In our view, editors are responsible for ensuring that diverse voices are present not only in reviewer selection but also on editorial boards and teams (cf. Lachert & Malik 2020). This also means keeping an eye on equity, such that BIPOC and LGTBQIA2S+ linguists are neither underrepresented nor overburdened with certain kinds of requests. We have heard concerns about bias in process and procedure (see now Daley 2021) but did not probe these concerns in this initial study. As the LSA Statement on Race (Linguistic Society of America 2019b) and the work of groups such as the Committee on LGBTQ+ [Z] Issues in Linguistics and Natives4Linguistics underscore, these issues cry out for work. For example, this recently appeared on Twitter from Katrina Daly Thompson (@putawaytheglobe, May 20, 2021): ‘Are there any academic journals that include authors’ pronouns in the byline? As an author, would you include yours if given the option? (Asking as a journal editor.’.

We also did not engage editors on questions about ethical practices. The LSA’s ethics statement (Linguistic Society of America 2019a) outlines guiding principles to support ethical choices in the range of settings in which linguists work. Relevant to our discussion here, the statement explicitly covers behavioral responsibilities (§2), responsibilities to communities and the public (§3), and responsibilities within the discipline (§4). This final section references such issues as maintaining confidentiality, respectful interchange of ideas, and fairness in peer review. Most linguistics journals are members of the Committee on Publication Ethics (COPE), which provides leadership in matters relating to professional integrity in scholarly research dissemination. COPE’s ethical guidelines for peer review (COPE Council 2017) foreground trust and the need for all participants to behave responsibly and ethically, noting at the same time that some reviewers may not be aware of their ethical obligations. Chief among these are anonymity and confidentiality. Anonymity is a principle of peer review because it is seen as essential for avoiding bias (e.g. Shatz 2004, Suárez et al. 2012). This is part of why reviewers should not seek information about authors when review is double-anonymized. Moreover, manuscripts are provided to reviewers for the sole purpose of peer review. They should not be used for any other purpose, and should certainly not be shared with others, cited in ongoing work, or discussed (most especially on social media) without the permission of both the editor and the author (which entails mutual agreement to set aside anonymity). Co-reviewing with a mentee—an important professionalization opportunity—requires permission before forwarding a manuscript to the student.

Overall, the responses to the survey and related discussions have shown us how much there is to do beyond this small first step, in terms of assessing where we are as a field, what best practices should be, and how we can shape field-wide standards.

4. Peer review in the time of COVID-19. By the time we began to analyze the survey responses, the world had been disrupted in ways we could not have imagined by the COVID-19 pandemic. Even beyond linguists who have suffered COVID-19, the pan-
demic had tangible and immediate impacts on academic publishing. In August 2020, we sent a query about the effects of the pandemic to the editors’ list, followed by another in early February 2021. Part of the motivation came from comments by Claire Bowern, then incoming editor of *Diachronica*, and various concerns raised on social media and in correspondence about the differential impacts of the pandemic, especially in terms of gender, early career researchers, and BIPOC scholars. As elsewhere, there has been robust discussion across academia about this topic (e.g. Cui et al. 2020, Oleschuk 2020, Viglione 2020, Watchorn & Smith 2020a,b). Drawing data from a broad survey conducted in May and October 2020, Watchorn and Smith (2020a,b) confirm stark pandemic-related effects, experienced differentially across field, career stage, and (predictable) social factors. For example, 37% of mid-career academic women (five to sixteen years post-Ph.D.) report severe impacts on their mental health and overall well-being, 51% of mid-career academics report being “severely limited” from conducting research (vs. 16% of late-career scholars), and 74% of humanities scholars feel “hampered in one way or another”. Linguistics encompasses the humanities, the social sciences, and the natural sciences, suggesting likely differences by subfield.

Our queries to the editors’ list brought widely divergent responses, and we do not have enough data to make generalizations. A few reported that the effects of COVID-19 have been somewhat marginal, writing, for example, that it “does not seem to have significantly impacted our business other than some authors and reviewers requesting additional time”. More report that reviewers are more frequently declining invitations, such as this comment from February 2021, which closely matched another response:

As an editor for two journals, getting reviewers has become incredibly challenging. I’ve also had people agree and then just never deliver. I’m hearing similar [stories] from others, sometimes going through 6–7 attempts to find a reviewer. So, [it is] also very time intensive.

Writing about what has happened with one linguistics journal, Toth (2020) surveyed reasons for declining to review, showing that in February and March of 2020, early in the pandemic, ‘37% of all reasons provided for not reviewing for the journal were connected to the social effects of the COVID-19 outbreak’.

As we write this, almost exactly a year after the pandemic began, it is clear that the situation has been dynamic over time, from initial shock and disruption to adjustment and development of new routines. Bowern 2020, one of several editorials in linguistics journals about the pandemic, documents an initial sharp increase in submissions from May to July at *Diachronica*, with things then returning to more normal numbers. At the Journal of English Linguistics, submissions continue to be disproportionately high over prepandemic rates, placing increased demands on reviewers and the editorial team alike. Submissions by women, to the extent this can be accurately judged, appear to remain at levels from previous years, though there are indications from other fields that gender is a key determinant of pandemic publishing: rates of prepublication and new research by women have either slowed or fallen behind those of their peers who are men (e.g. Oleschuk 2020, Viglione 2020). Many linguistics journals are facing greater difficulty in finding reviewers, and time to decision has doubled in some cases. A report shared on social media (Facebook, Twitter) in early February 2021 by Johan Rooryck, editor of *Glossa*, shows a 22% increase in time to decision in 2020 over 2019 figures. Submissions themselves were up by 19%, while the number of reports received decreased by 2% overall, despite the higher number of reviewer invitations in response to higher submission rates. These figures underscore the immediate repercussions of COVID-19 on linguistics journal publishing.
Still, as one editor points out in correspondence (and see also Bowern 2020), the biggest effect should come in future years. We have a pile of papers to publish so most papers people worked on during the pandemic won’t appear until afterwards, and likewise, papers that could have been written but weren’t because of pandemic obligations weighing scholars down won’t be noticed as absent until a year or two down the road.

And we have been advocating within our own institutions for allowances in processes related to reappointment, tenure, promotion, and salary adjustment that take the long tail of the pandemic on productivity in all its guises into consideration. This will be particularly important for research that relies on in-person, face-to-face collaborative fieldwork.

Overall, responses we received and the conversations we have had about the impacts of COVID-19 suggest that, like in so many other ways, the pandemic has exposed and exacerbated already-present underlying structural issues: difficulties in finding reviewers, challenges of publishing for early career researchers, and imbalances deriving from equity, diversity, and inclusion concerns were always there, but have been made more pressing and more visible with the pandemic.

5. Conclusion. This paper has tried to crack open what Schaffalitzky de Muckadell and Nedenskov Petersen (2017) call, in a different context, the ‘black box of journal editing’ for linguistics. Writing this paper has changed our own grasp of peer review, and our aim is to pass that on and engage the broad population of people who submit to, review for, edit, and ultimately read linguistics journals, to see both how the process works and how it does not, in the hopes of improving it. Our results indicate that respondents work hard to set the best practices they can. Survey results and discussions and correspondence with editors and others have deepened our commitment to peer review, but with that comes a sharper awareness of how it needs to be, and can be, improved. One editor reflected on peer review, noting:

My feeling … is that the value of the review process isn’t understood by many people in the field and there isn’t a culture of seeing it as a worthwhile and necessary practice. Authors see it just as a hurdle to get over (rather than an opportunity to measure up their work) and people asked to review see it as a chore rather than as service.

This engagement is what we need to foster (and see Curtin et al. 2018).

In the meantime, all involved can play a part in developing, ensuring, and improving best practices. We see clear needs for more and better mentoring across the board, from how to write a review to how to navigate the submission and revision process, to becoming an editor. If the feedback we received is any indication, there is genuine desire within linguistics for consistent and informed training in academic publishing. For authors, this includes careful selection of venue, patience, and real care in revision. For reviewers, this begins with responding promptly to requests, especially if the invitation must be declined, and delivering reports on schedule that focus on issues of substance, that are constructive, and that aim for improvements that honor the author’s intentions, whether the review is critical or supportive. And we are equally responsible for developing standards that uphold ethical and equitable practices, not only in peer review but also in all aspects of publishing and associated professional and professionalization activities.

14 We may even fit into a taxonomy like this one: http://deevybee.blogspot.com/2010/09/science-journal-editors-taxonomy.html; many linguistics editors aim to be paragons. We tip our hat to the respondent who called our attention to this.
Editors are clearly eager for better and more innovative approaches to peer review. We have a lot of work to do: building better collaborative editorial teams, bringing in new and innovative perspectives, diversifying the voices that contribute to scholarship, and improving time to decision. Our conceptions of ethics, equity, and inclusion are advancing, but this survey underscores the need for more effort on those fronts. Clearly, some problems will not be easily overcome, such as the behavior of real-world Reviewer 2 (§3.3) who resides in small ways in all of us. The challenge is to keep those tendencies at bay while striving to be constructive. There is enough common ground in current practices that we see potential for explicit field-wide standards, an issue that seems suited for the Committee of Editors of Linguistics Journals. As noted in §2, this group is not restricted to the membership of the LSA; its reach (disciplinary, tradition-based, geographic) is broad.

Reviewers, editors, and authors play different roles, obviously, and many of us play them all at different points. Whether or not peer review operates at maximal capacity is not simply about how easy it is, or how much of a meritocracy it is, or how much the numbers suggest it works or not. Rather, it is about whether the end product is research that better reaches its target audience and advances the field. We again take the firm stance that papers get better, not worse, through the process. This is the service element noted in the quote above, and editors carry the responsibility for ensuring that is the case, with concrete help from reviewers and authors. It is a team effort, in the end, one that, as our survey indicates, is facilitated by kindness, generosity, and grace (Figure 4; @acaguy, January 17, 2019).

One theme across many reactions to early drafts of this paper was how interconnected the roles of editors, authors, and reviewers are and how grasping that interconnectedness is critical for early career researchers. Knowing norms and processes in our communities helps newcomers not only to engage better with the reviews they receive and thus produce better revisions, but also to write better reviews as they in turn are on-boarded into that crucial role. At the same time, there is an opportunity for experienced reviewers to reflect critically and openly on their own approach to peer review. Of course, knowing more about the practice of peer review should help authors at all career stages, both within the academy and outside it, navigate the process better, from selecting the right venue to interpreting editorial decisions and engaging successfully with revisions.

We have aimed to provide some clarity about what peer review looks like from the inside. With luck, we have contextualized how things are experienced on the outside.

15 There is also need for authors to be generous. An author may not agree with or understand all points raised by reviewers. This does not necessarily mean that Reviewer 2 has made an appearance—only that people have different perspectives. This is part of the richness of academic discussion, and relates to our points above (§3.5) regarding fulsome and thoughtful engagement with reviews when revising.
We hope this paper contributes to building sound and transparent processes that can help us all do better science for our field, the communities with whom we collaborate, and the public. To make the progress we need ultimately demands that the whole community know the peer review process and engage actively in understanding our various roles, including working to execute them better.

Every linguist needs to understand peer review, engage with colleagues and students about it, and participate in improving the process. We hope that this is the beginning of a longer discussion in the field, among authors, reviewers, and editors.

**APPENDIX: GLOSSARY OF PUBLISHING TERMINOLOGY**

**ACCEPT**: The journal will publish the paper and no further revisions are required. At this stage, the manuscript is forwarded to the copyediting queue. The outcome is final (unless an academic integrity violation is uncovered, such as plagiarism, evidence of publication elsewhere, data falsification, etc.).

—*(ACCEPT with) MINOR revision*: The journal has accepted the paper conditionally, on the expectation that the author continue to revise in accordance with feedback from reviewers and the editor. The required revisions are not substantive. The paper has yet to be accepted definitively.

—*(ACCEPT with) MAJOR revision*: See Revise and resubmit.

**ANONYMOUS REVIEW**: See Review.

**COPYEDITING**: The process of revising an article, prior to submission to the publisher’s production department, to improve readability and to check grammar, spelling, punctuation, and other elements of style in accordance with the journal’s style sheet and practices. This is also the stage where editors and editorial assistants check for mistakes, inconsistencies (e.g. terminology, hyphenation), and repetition of content. Many editors will return copyedits to authors to follow up on questions they have or to confirm that the suggested changes are acceptable.

**DESK-REJECT**: The journal has decided to not publish the paper, and no peer review was sought. The paper cannot be resubmitted to the same journal. The outcome is final, unless the decision explicitly indicates that resubmission is possible assuming specific changes to the manuscript.

**EDITOR (IN CHIEF OR EXECUTIVE)**: The editor is the journal leader, ultimately responsible for decisions, reviewer selection, and feedback to authors. They liaise with other members of the editorial team and the publisher, including the production team that oversees publication in paper and digital format. They are expected to be researchers and to have expertise in the field in which the journal publishes. They are responsible for the reputation and the standards of the journal.

**ASSISTANT EDITOR**: See Editorial Assistant.

**ASSOCIATE EDITOR**: An editor with duties delegated from the Editor (in Chief or Executive). Typically has responsibility with regard to article decisions and may be assigned to a thematic subset of journal submissions.

**EDITORIAL ASSISTANT**: Provides administrative support to the editor, performing preliminary checks of submissions for anonymization, content completeness, legibility of tables and figures, and the like. They also track submissions and are the primary point of contact for authors who reach out the journal. For some journals, the editorial assistant is responsible for the bulk of copyediting and checking proofs before passing articles on to editors for final approval, prior to submission to the publisher’s production team.

**EDITORIAL (ADVISORY) BOARD**: A consulting body for the editorial team, consisting of subject-matter experts representing the scope of the journal. Depending on the journal, board members may be active reviewers, or their positions may be honorific.

**PROOFS**: A preliminary version of the article that shows how it will look in a print or online publication. They are intended to provide authors and editors an opportunity to perform a final check of the article, prior to publication. Revisions at this stage are limited to stylistic corrections, typos, and minor wording changes; substantive change is not possible.

**REJECT**: After peer review, the journal has decided to not publish the paper. Once a journal rejects a paper, it cannot be resubmitted to the same journal for consideration, regardless of how much revision it has undergone. The outcome is final.

**REVIEW**: The quality assurance process of having research vetted by subject-matter experts.

—**DOUBLE-ANONYMOUS REVIEW**: The identities of both the reviewers and the author are unknown to each other. This was traditionally referred to as ‘double-blind’ review, a term avoided here as ableist.

—**SINGLE-ANONYMOUS REVIEW**: The identities of the reviewers are unknown to the author, but the identity of the author is known to reviewers. This was traditionally referred to as ‘single-blind’ review, a term avoided here as ableist.
—**Triple-anonymous review:** The identities of reviewers and authors are unknown to each other and to the editor handling the manuscript. This has been referred to as ‘triple-blind’ review, a term avoided here as ableist.

**Reviewer:** A subject-matter expert sought out by editors to provide an assessment of the merit of a journal submission, with an emphasis on descriptive, empirical, and theoretical adequacy.

**Reviewer 2:** Reviewers are assigned a number that reflects, depending on the journal, the order in which they were invited to review, the order in which they accepted the invitation to review, or the order in which they submitted their review. This numbering system protects reviewer anonymity. Reviewer 2 has become a trope in academic publishing, embodying all of the habits of bad peer review. There is no correlation, however, between reviewer number and their recommendation to the editor (e.g. Weisberg 2017). There are ways in which nobody is Reviewer 2 and ways in which Reviewer 2 resides in small ways in all of us. Web searches for ‘Reviewer 2 joke’ or ‘meme’ may offer a chance to laugh about this topic.

**Revise and Resubmit (R&R):** Sometimes also referred to as a major revision, depending on the journal’s (online) system. This decision indicates that the editor and reviewers see value in the paper but that it does not yet meet the journal’s publication standards and substantial revision is required. The paper is neither accepted nor rejected. Rather, the editor invites a revised submission, and the paper will again be sent out for peer review. There is no guarantee that the paper will ultimately be accepted, but a revise and resubmit maintains the possibility that it may be. In our experience, this is the most common decision on initial submissions.

**REFERENCES**


Helmer, Markus; Manuel Schotttdorf; Andreas Neef; and Demian Battaglia. 2017. Gender bias in scholarly peer review. eLife 6:e21718. DOI: 10.7554/eLife.21718.


[adarcy@uvic.ca] [Received 15 March 2021; revision invited 13 May 2021; revision received 2 June 2021; accepted 7 June 2021]