Research Methodology

Why Do Articles Get Rejected by International Journals ?

Richard Watson Todd
King Mongkut's University of Technology Thonburi

Abstract

With increasing pressure for university staff and students to publish in international refereed journals, researchers will benefit from guidance on how decisions to accept or reject articles are made by these journals. This paper examines 28 reviewer reports in applied linguistics, and categorises the 115 individual criticisms made in two ways. First, the criticisms were categorised based on the article section that prompts the comment, and second, they were categorised based on the research quality criterion that the comment focuses on. There is no clear indication that any particular section or criterion is most likely to lead to an article being rejected. Rather, rejection typically appears to be based on an accumulation of comments of different types. Example comments linked to guidelines for writing articles are given to help novice researchers to produce articles less likely to be rejected.

1. Introduction

Universities are increasingly requiring their academic staff and postgraduate students to publish research in international refereed journals (Cheng, 2006; Huang, 2010). For many staff and students in applied linguistics, this is a major challenge which appears to have little chance of success since many of the major journals have acceptance rates of 20% or lower (Egbert, 2007) and the staff and students are competing with experienced researchers for the few slots for publication available. In this paper, I will examine some of the reasons why journals reject submissions in the hope that this will provide guidance for prospective authors.

Most of the existing advice for prospective authors of research articles is derived from genre analyses and thus focuses on the generic or linguistic features of articles (e.g. Swales, 2004; Swales and Feak, 2004). While undoubtedly useful for authors, articles are rarely rejected on generic or linguistic grounds (Jaroongkhongdach et al., 2012). Rather, most rejections are based on content or technical research issues (Gosden, 2003; Mungra and Webber, 2010).

To identify what sorts of content or research issues are most likely to lead to rejection, we need to examine reviewers' comments. The typical process facing an article submitted to a reputable international journal starts with an in-house evaluation of the article to see if it is of sufficient quality to send to reviewers. While in-house rejection rates are reputed to be increasing (Zuengler and Carroll, 2010), this stage of the process is not open to examination. Those

articles deemed to be of sufficient quality and suitability by the editors are sent to reviewers for comments and recommendations (typically, requiring minor revisions, requiring major revisions, or rejection). Those papers requiring revisions may go through this process two or three times as the drafts of the article change. Although still an "occluded genre" (Gosden, 2003, p. 87), reviewers' comments are to a limited extent available for analysis. This paper examines 28 referee reports to attempt to identify points perceived as problematic in articles submitted for publication in international refereed journals.

2. Methodology

2.1 The data

The 28 reviewer reports examined in this paper are ones I have written for five different international journals over the last few years. With all the reports being from the same reviewer, there may be some limitations in the topics focused on, with some being under-represented and others over-represented, but the difficulties of gaining access to reviewer reports mean that obtaining a fully representative sample is impossible. Of the 28 reviews, three recommend minor revisions and thus the submissions are likely to be accepted, five recommend major revisions with the decision for acceptance dependent on the revised version, and 20 recommend rejection. These rates reflect the overall acceptance rates of the journals, suggesting that the reports are fairly typical. All of the reports are included in the analysis, since even the criticisms made on those articles likely to be accepted can shed light on the content and research issues underpinning journals' decisions on articles. A typical report is 500-800 words long and consists of a general introductory paragraph which includes the recommendation followed by 5-20 detailed criticisms identifying points needing revision. Some of these criticisms are minor (e.g. linguistic infelicities that suggest the need for better proofreading such as "who did the dog barked at" and "to same students") which would not be a cause for rejection, although they could stimulate an unconscious bias in the mind of the reviewer who would then tend to view other aspects of the article more negatively. These minor criticisms are straightforward to correct and are not included in the analysis. In total, 115 points of criticism are analysed.

2.2 Data analysis

The 115 individual criticisms were first grouped to identify similar criticisms with different wordings. For instance, "The conclusion and discussion sections appear to simply repeat the findings without adding any insights or generalisations and thus still leave me wondering about the value of the research", "A large part of the Discussion simply repeats the findings", and "The discussion just presents a summary of the findings" were all summarised as *Discussion is a repetition of the findings*. This process resulted in 51 grouped criticisms.

These grouped criticisms were then categorised according to two sets of criteria:

1. *Article section:* introduction and literature review, purpose (e.g. research question), methodology, results, discussion, other (e.g. title, abstract).

These sections follow the standard format for research articles, and the categorisation involves identifying the section which is primary in prompting the criticism. All grouped criticisms were categorised into article section. Categorisation into article section may allow sections particularly likely to prompt criticisms, and thus those which are particularly problematic, to be identified. In addition to the six article sections, a further category of 'overall suitability' was created for those criticisms which concerned the whole article rather than being linked to a section.

2. *Research quality criteria* (based on Jaroongkhongdach et al., 2012): justification, clarity, coherence, appropriacy, awareness.

These criteria allow the criticisms to be categorised on the basis of their contents (rather than what prompts them, as in categorisation based on article section). Jaroongkhongdach et al. (2012, p. 197) define these five research quality criteria as follows:

- *Justification:* "reasoning provided for decisions made in research"
- *Clarity:* "the sufficiency of descriptions or explanations of a term/concept/procedure, and the style of writing that makes the term/concept/procedure easy for an intelligent general reader to understand"
- *Coherence:* "the logical relationships within a section or across sections in terms of contents or ideas"
- *Appropriacy:* "the match/compatibility between two or more potentially related components"
- *Awareness:* "the thoughtful concern of alternative views or of possible impacts of research decisions"

42 of the 51 grouped criticisms (accounting for 102 individual criticisms) could be categorised by research quality criteria. Although not providing full coverage of the criticisms, as with categorisation into article sections, this may allow particularly problematic aspects of research to be identified. Criticisms not falling into these categories include *Article is too long and Literature is dated.*

The frequencies for each of these criteria were counted. Key grouped criticisms which appeared to be particularly important in decisions for rejecting articles (either because they appeared frequently across articles or because they were identified as the key issue leading to rejection) were identified. Details of the actual contents of the reviewer reports for these points are shown to provide a more in-depth perspective on reasons for rejecting articles.

3. Results

3.1 Grouped criticisms by article section

Table 1 presents the frequencies of individual criticisms and grouped criticisms by article section to identify those sections which may be particularly likely to lead to a submission being rejected.

Those grouped criticisms consisting of three or more individual criticisms are shown to concretise the data (with number of individual criticisms in each group given in brackets).

Table 1 Grouped criticisms by article section

Article section	No. of grouped criticisms (N = 51)	No. of individual criticisms (N = 115)	Sample grouped criticisms
Introduction and literature review	8 (15.7%)	19 (16.5%)	No link between literature and research (7) List-like with no argumentation (3) Missing key references or aspects requiring discussion (3)
Purpose	4 (7.8%)	8 (7.0%)	No reason why research is useful (3) Lack of clarity in purpose (3)
Methodology	15 (29.4%)	31 (27.0%)	Lack of details (10) Claimed principles not followed (3) No reporting of rationales for decisions (3) Unclear foundation or sources of data in general (3)
Results	9 (17.6%)	23 (20.0%)	Poor interpretations (7) Inappropriate statistics (3) Very limited data presented (3)
Discussion	8 (15.7%)	22 (19.1%)	Dubious explanation of results or unsubstantiated claims (6) Discussion not based on findings or not related to purpose (4) Discussion is repetition of literature review (3) Discussion is repetition of findings (3)
Other (e.g. title)	4 (7.8%)	5 (4.3%)	
Overall suitability	3 (5.9%)	7 (6.1%)	Inappropriate topic for journal (3) Article is too long (3)

From Table 1, we can see that the methodology section of articles is the one stimulating most criticism with *Lack of details concerning the methodology* being the most frequent grouped criticism. However, the literature review, results and discussion sections also prompt criticism, suggesting that no single section of articles can be identified as the one most likely to lead to article rejection.

3.2 Grouped criticisms by research quality criteria

The five criteria of research quality suggested by Jaroongkhongdach et al. (2012) were also used to categorise the criticisms. Not all criticisms fell into these categories (ones not falling into these categories are not considered in this section), and Table 2 presents the frequencies of criticisms by these criteria and the most frequent grouped criticisms.

Table 2 Grouped criticisms by research quality criteria

Research quality criteria	No. of grouped criticisms (N = 42)	No. of individual criticisms (N = 102)	Sample grouped criticisms
Justification	7 (16.7%)	13 (12.7%)	No reason why research is useful (3) Missing key references or aspects requiring discussion (3)
Clarity	14 (33.3%)	35 (34.3%)	Lack of details in methodology (10) Lack of clarity in purpose (3) Unclear foundation or sources of data in general (3) Very limited data presented (3)
Coherence	9 (21.4%)	19 (18.6%)	No link between literature and research (7) Claimed principles not followed (3)
Appropriacy	7 (16.7%)	22 (21.6%)	Discussion not based on findings or not related to purpose (4) Poor interpretations (7) Inappropriate statistics (3)
Awareness	5 (11.9%)	13 (12.7%)	Dubious explanation of results or unsubstantiated claims (6) No reporting of rationales for decisions (3)

From Table 2, clarity is the criterion generating most criticisms, but the frequency of all of the criteria is enough to suggest that no single criterion can be identified as the most likely cause for rejection.

3.3 Key criticisms

Tables 1 and 2 provide the frequencies of the various criticisms, but it is not clear whether these frequencies should be linked to causes for rejecting submissions. Indeed, the most frequent grouped criticism, namely, *Lack of details concerning the methodology*, is relatively straightforward for authors to revise since additional details can be added easily. Indeed, this criticism appears to be linked to article rejection in only one case, where the lack of details was so pervasive that it was impossible to understand how the research had been conducted. Nevertheless, this criticism

is still important since a recommendation for rejecting an article may be due to the cumulative effects of numerous minor issues, rather than a single main issue. The frequency of criticisms of missing methodological details suggests that, as this is an area which is relatively straightforward to revise, *Lack of details concerning the methodology* is a point prospective authors should be aware of so that the probability of an article being rejected because of the accumulation of minor criticisms is reduced. In contrast, some other criticisms are not frequent but are the main cause for rejection, such as *Inappropriate topic for journal and Invalid poor quality main instrument* (neither of which are discussed further below since they are self-evident). In this section, I will provide the verbatim comments (with some amendments to preserve anonymity where necessary) for those eight grouped criticisms that are either very frequent (leading to the possibility of a cumulative rejection) or very salient (being the main issue in rejecting an article). I hope that providing verbatim criticisms from reviewer reports can allow researchers new to international journal publication to become familiar with the kinds of comments typically given on articles. I also hope to be able to induce some guidelines that may help researchers to be able to write articles that are less likely to be rejected.

1) *No link between literature and research*

Criticism 1.1:

"There is a general lack of foundation to the article. This may be related to how the previous literature in the area is used. For example, the author lists previous research into SMS without showing how the current study builds on these. The lack of a clear foundation means that the interesting findings of the study are not discussed in an insightful way (indeed, the discussion section of the article is not really related to the findings)."

Criticism 1.2:

"The rationale for conducting the research is unclear. The introduction presents the teaching context and then jumps to the purpose of the study without providing any links between the two. Similarly, while the literature review provides a useful overview of feedback on students' writing, it does not lead to the purposes of conducting the research."

Guideline: The term 'literature review' is perhaps a misnomer, as the purpose of this section is to give an argument providing the rationale for the research, rather than simply provide a review of the existing literature. Perhaps because of this misunderstanding, some novice researchers use the literature review to list (with some details) the previous key studies in the area of their research. Instead, although there may be paragraphs providing background knowledge on areas readers are unlikely to be familiar with, the majority of the literature needs to be presented in a way that builds an argument for why the research is being conducted.

2) *No reason why research is useful*

Criticism 2.1:

"Most seriously, I am not clear how the research adds to the existing literature on L1/L2

use. Generally, in a given field of research (such as language of classroom communication), there is a longitudinal move from initial descriptions towards explanations and evaluations. The current article is descriptive in nature, somewhat unexpected for a field of research with a history of over 30 years. In itself, this might not be a problem, but there is no clear indication of how the article adds to the field."

Criticism 2.2:
"I am not really clear what the point of the article is. On a simplistic level, the article shows that students interact in different ways for different task types – a finding which is not very surprising. The literature review, while well-argued and covering a good range, appears to state that there has been a lot of previous research in the area of floor management, including the floor structure of different tasks. With such a wealth of existing literature, it is unclear how the current article adds anything."

Criticism 2.3:
"The purpose of the research is not altogether clear. While the literature review covers a fair amount of ground, it does not lead into the present study. No research questions are given and, by the end of the article, I'm not really sure why I should be interested in the study."

Guideline: The author needs to persuade the audience that the research is worthwhile since it serves a valid and useful purpose. The purpose can be either theoretical (e.g. challenging an existing theory) or practical (e.g. having implications that could change current practice). Often, the purpose is presented as filling a gap in the literature which is argued to exist in the literature review. However, not all gaps that exist are worth filling (e.g. I am not aware of any research into the effects of the teacher's sock colour on learning), and many are of dubious value (e.g. implementing a standard questionnaire in a new context). The research therefore needs to show a worthwhile purpose argued for in the literature review, stated clearly (usually in research questions) and highlighted in the discussion.

3) *Lack of details concerning the methodology*
Criticism 3.1:
"A lack of rigour in the article is also a concern. This is perhaps best illustrated in the error analysis. How was this conducted? How were errors identified and categorised? There is no need for great detail concerning such points, but some statements of principles used would be useful. Similarly, a lack of details about the data analysis in the survey leaves me wondering why the percentages reported don't add up to 100."

Criticism 3.2:
"There are some problems with the methodology of the study. Given the centrality of the interview and questionnaire to the research, I would have expected to see what questions were asked. This is especially important for the questionnaire as it is impossible to evaluate the findings. Without knowing the questions and the meaning of the rating scales, I don't

know what, say, a mean of 3.66 for instrumental orientation means."

Guideline: Enough details are needed to allow readers to understand the findings, make judgments about their validity and reliability, and make their own interpretations. Word count permitting, too many details are better than too few.

4) *Claimed principles not followed*

Criticism 4.1:

"More seriously, the research claims to 'develop an insider's view of the phenomenon', but the approach of counting frequencies of reference and providing illustrative quotations means that this is not achieved."

Criticism 4.2:

"In the methodology section, the author states that the data will be analysed using grounded theory analysis, but there is no evidence concerning how coding was conducted or how themes were identified. Indeed, from the findings, it appears that the identified themes follow the interview questions, and thus that no attempt to apply a grounded theory analysis was made."

Criticism 4.3:

"On p. 7, the research design is termed "A mixed-methods research design", but no qualitative data is used."

Guideline: It may be tempting to claim that a certain methodological approach has been used in the research since this appears to provide a veneer of credibility. However, most methodological approaches have clear implications for, and even restrictions on, how the research needs to be conducted. If the approach is not implemented as intended, then claiming to follow it will lead to an adverse reaction in the reviewer that clearly outweighs any benefit from making the claim.

5) *Poor interpretations*

Criticism 5.1:

"Some of the interpretations of the findings are not very persuasive. For instance, on p. 23 the authors interpret the findings as suggesting that students prefer lectures; yet, on p. 24 the students appear to complain about lectures and the lack of opportunities to practise."

Criticism 5.2:

"The findings are very weak. Quotations, where included, are generally of clause length and thus do not provide the necessary context for the reader to interpret the quote. More seriously, many points are made without any reference to the data (even in the form of very short quotes). For instance, in the final paragraph of 4.1 concerning school commitment, no evidence concerning the school under investigation is provided. The paragraph reads more like a paragraph from a literature review than part of the findings. Furthermore, where evidence is provided, often this does not match the arguments being made. In the first

paragraph of 4.1, for example, it is argued that the CALL program encourages learning and boosts confidence, but the quotations concerning these points concern enjoyment."

Guideline: Interpretations generally aim to highlight key aspects of the findings for the readers, often by focusing attention on a particularly salient finding or by summarising a pattern in the findings. Interpretations should concern those aspects that are truly worth highlighting (and not, for example, simply be a repetition of a table in prose) and should be valid and reasonable. To allow readers to judge the validity of the interpretations, findings should be given in enough depth.

6) *Discussion not based on findings or not related to purpose*

Criticism 6.1:

"Several contributions of the research are claimed in the discussion section, but some of these are already well accepted in the field, while others are not based on the data. For instance, there is discussion of the motivations of teachers (bottom p. 17) and reasons for variability between teachers (top p. 19), yet there is no data concerning motivation or reasons."

Criticism 6.2:

"The implications are not derived from the findings. For instance, there is nothing in the findings asking that the contents of programs be regularly updated. The conclusion is also not linked to the findings (e.g. there is nothing in the findings to indicate a problem of access that could be solved by increasing the number of licences). Indeed, at one point, the conclusion conflicts with the findings: in point 2 of the conclusion it is stated that monitoring is limited to the teachers' resourcefulness, whereas on p. 12 it is stated that the program includes a tool to monitor progress."

Guideline: Although the researcher may be tempted to use the discussion section as a platform for arguing for their beliefs, the discussion must be (at least initially) based on the findings which take priority. Care should be taken that arguments which are not derivable from the findings are not presented.

7) *Discussion is repetition of literature review*

Criticism 7.1:

"Much of the discussion section appears to be a repetition of the literature review, rather than a true discussion of the findings. As an alternative the author could be suggesting things like certain plateaus of proficiency that need to be reached before improvements in the comprehension of different text types become apparent."

Criticism 7.2:

"The conclusion shifts the focus away from instructional strategies and back to the general benefits of CMC. It thus appears to be more like a reiteration of the literature review than

a conclusion to the article."

Guideline: The discussion section is sometimes (with good reason) titled 'Discussion of the findings' and should use the key findings as the starting point for discussion. Where relevant, the discussion can refer back to the literature review, but it should go beyond what has already been stated in the literature review by, for example, focusing on salient findings which have implications contradicting or not previously considered in the literature.

8) Dubious explanation of results, or unsubstantiated claims

Criticism 8.1:

"The discussion and conclusion have the feel of shoehorning the data into a predetermined explanation. While norms are one way of explaining the data, there is no direct evidence for norms, and several other explanations are possible. Furthermore, although the addition of qualitative data from another paper does shed light on the findings, this additional data appears to be used where arguments for norms based solely on the quantitative data are not convincing. The overemphasis of interpreting data as being due to norms also appears earlier where it is claimed that any similarities between classes are due to institutional norms (again, several alternative explanations, such as the homogeneity of the students, are possible)."

Criticism 8.2:

"The discussion claims that students have a 'confused mindset' (a doubtful description in itself), but also that their opinions can be 'explained in a rational manner'. Much of this explanation is of a very dubious post hoc nature. For example, the preference for NSTs for both speaking and reading is explained based on the high correlation between phonological awareness and reading ability. If this dubious claim is correct, I don't see why it wouldn't also explain a preference for NSTs for listening. Nearly any combination of results can be 'explained' on doubtful bases such as this, but explanations based on other literature from the same area (such as the differences between explicit and implicit beliefs in students' preferences for NSTs or NNSTs) is more persuasive. It would also be useful if the results from this study were compared to results from similar studies."

Guideline: Researchers have their own reasons and beliefs which motivate them to conduct the research. There is often a tendency for discussion to centre around such beliefs. However, a discussion which considers the findings from several perspectives or which presents several competing explanations for findings is more persuasive than one which presents a single viewpoint based on the researcher's existing beliefs.

4. Discussion

The findings suggest that no single article section or research quality criterion is particularly likely to lead to an article being rejected, since the articles receive on average over four substantive

comments each. Reading the introductory paragraphs to the reports, only three articles were clearly rejected on the basis of a single comment (two for *Inappropriate topic for journal, and one for Poor quality main instrument).* Rather, most articles were rejected based on an accumulation of criticisms. The following extract from an introductory paragraph is typical of the explanation of reasons for rejection:

> *"The article under review has some potential in this area, but its purely descriptive nature (rather than being explanatory) means that it is unclear how it contributes to the field. Together with several weaknesses in the methodology and the writing up (detailed below), this means that the article cannot be accepted in its present form."*

Some of these accumulating criticisms are relatively frequent but also fairly straightforward to revise, such as *Lack of details concerning the methodology*, the frequency of which led to the overall high frequencies for the methodology section and the criterion of clarity. By themselves, these criticisms have little effect on a decision to reject an article, but when combined with four or five or more other criticisms, they may lead the reviewer to pass a tipping point and decide to reject the article. Being aware of the types of criticisms which are frequently made by reviewers may enable researchers to reduce the number of accumulated minor criticisms and thus lessen the chance of their article being rejected.

Jaroongkhongdach et al. (2012) identified five criteria of research quality based on the literature analysing research articles and from an analysis of the articles themselves. These five criteria also seem applicable to reviewer comments, with over three-quarters of the individual criticisms assignable to a quality criterion. Both Jaroongkhongdach et al. and this paper examine research in applied linguistics, a soft-applied discipline where variables are varied, causal connections are tenuous, synthetic inquiry strategies are used, authors typically make personal stands in the text, and the readers' need to be involved in the negotiation of knowledge-making is acknowledged (Hyland, 1998). These characteristics of applied linguistics mean that most research in the field aims to persuade readers through the discourse of the article, rather than simply presenting relevant facts within the context of the discipline as the main persuasive strategy as in the hard-pure sciences. A prerequisite for persuasive discourse is comprehensibility, a characteristic to which both clarity and coherence contribute. In addition, a well-argued text with valid justifications and awareness of problems is generally more persuasive than one without these features (Cockcroft and Cockcroft, 2005). The research quality criteria therefore may be subsumed into the broader purpose of persuading the reader of the value and usefulness of the research. Reviewers' comments leading to rejection, then, indicate that the article has not persuaded them of the value of publishing the research.

I hope that this article may have raised the awareness of researchers, especially those new to publishing in international refereed journals, of what to expect from reviewers and that paying attention to the points covered in the article may reduce the chances of research being rejected.

Finally, I also hope that this article avoids most of the problems covered in reviewer comments that can lead to rejection.

References

Cheng, A. (2006). Analyzing and enacting academic criticism: The case of an L2 graduate learner of academic writing. *Journal of Second Language Writing*, *15*, 279-306.

Cockcroft, R. and Cockcroft, S. (2005). *Persuading people: An introduction to rhetoric*, (2nd ed.). Basingstoke, England: Palgrave Macmillan.

Egbert, J. (2007). Quality analysis of journals in TESOL and applied linguistics. *TESOL Quarterly*, *41*, 157-171.

Gosden, H. (2003). 'Why not give us the full story?': Functions of referees' comments in peer reviews of scientific research papers. *Journal of English for Academic Purposes*, *2*, 87-101.

Huang, J. C. (2010). Publishing and learning writing for publication in English: Perspectives of NNES PhD students in science. *Journal of English for Academic Purposes*, *9*, 33-44.

Hyland, K. (1998). Persuasion and context: The pragmatics of academic metadiscourse. *Journal of Pragmatics*, *30*, 437-455.

Jaroongkhongdach, W., Watson Todd, R., Keyuravong, S. and Hall, D. (2012). Differences in quality between Thai and international research articles in ELT. *Journal of English for Academic Purposes*, *11*, 194-209.

Mungra, P. and Webber, P. (2010). Peer review process in medical research publications: Language and content comments. *English for Specific Purposes*, *29*, 43-53.

Swales, J. M. (2004). *Research genres*. Cambridge, England: Cambridge University Press.

Swales, J. M. and Feak, C. (2004). *Academic writing for graduate students: Essential tasks and skills*, (2nd ed.). Ann Arbor, MI: University of Michigan Press.

Zuengler, J. and Carroll, H. (2010). Reflections on the steady increase in submissions. *The Modern Language Journal*, 94, 637-638.

Author:

Richard Watson Todd is Associate Professor in Applied Linguistics at KMUTT. He holds a PhD from the University of Liverpool and is the author of numerous articles and books, most recently, *Much Ado about English* (Nicholas Brealey Publishing).
irictodd@kmutt.ac.th