A Publication Model for Open-Access Books in Linguistics

Stefan Müller and Martin Haspelmath

March 1, 2013

1 Allgemeine Angaben

Erstantrag auf Gewährung einer Sachbeihilfe für das Projekt: *A publication model for open-access books in linguistics*

1.1 Antragsteller

Name: Stefan Müller, Prof. Dr. Ing. habil.

Name: Martin Haspelmath, Prof. Dr. habil.

1.2 Thema

Language Science Press: A publication model for open-access books in linguistics

1.3 Förderprogramm / Fach- und Arbeitsrichtung

Ausschreibung: Wissenschaftliche Monographien und monographische Serien im Open Access
Fach: Sprachwissenschaft

1.4 Voraussichtliche Gesamtdauer

Das Vorhaben wurde noch nicht begonnen. Das Projekt ist auf die Dauer von 2 Jahren angelegt.

1.5 Antragszeitraum

Das Projekt ist auf die Dauer von 2 Jahren angelegt.
Gewünschter Beginn der Förderung: So früh wie möglich.

1.6 Zusammenfassung

The purpose of the project is to establish a publishing unit called *Language Science Press* for publishing linguistic monographs in several series. We want to build and maintain a community and community infrastructure that can be used for "Platinum" open-access Publications. Such a platform is seen as desideratum by the linguistic community as is evidenced by the large number of prominent linguistis who support our initiative and want to collaborate with us on implementing such a platform and create a high impact brand. The platform that we want to develop will run several linguistic series at minimal costs. In addition to reviewing that is done by the research community anyway we also try to distribute the work of typesetting and proofreading and solve the problem that these costly and time consuming parts of the production process pose.

While most of the software that is needed to produce linguistic monographs is available, some crucial parts are still missing. There is workflow software for producing books, for instance Open Monograph Press (OMP), which will be used in the project. LaTeX packages for typesetting exist. However, the problem of converting LaTeX documents into e-books and XML is not solved. We plan to establish a standard set of LaTeX macros that can be transformed into XML and the e-book formats. Furthermore we want to extend the OMP software by versioning and open review aspects.

The process of setting up Language Science Press will be accompanied and monitored by a business economist who will work on the documentation of the costs and ensure the development of a sustainable business model.

2 Ausgangslage / Stand der Forschung bzw. Entwicklung

2.1 Darstellung

2.1.1 State of the Art: book publication in linguistics

2.1.1.1 The book crisis

The publication of scientific books is in a serious crisis. The journal crisis in the natural sciences has been known for quite a while, but the same applies to books in fields which (like linguistics) make heavy use of them. The efficiency with which book content can be distributed electronically contrasts strikingly with the inefficiency of distributing printed books. With contemporary technology, it is clear that electronic scientific content should be seen as primary, with printed versions as possible derivatives.

Where e-books are available through the traditional publishers, they cost as much as printed books, so the much greater efficiency of electronic distribution has not translated into lower prices (cf. Müller, 2012). On the contrary, as commercial book publishers have to make greater profits in the wake of increasing market concentration, book prices have been rising much more than other prices. This is evidently due to market dysfunction: Authors have little incentive to prefer publishers with lower book prices to publishers with higher prices. What matters to book authors is the prestige of the publisher's imprint, because reputation by association with other high-quality work is an important ingredient in building scientific careers (Haspelmath, 2012).

As a result, books are still normally published in the traditional way, but this leads to paradoxical results, totally counter to the spirit of science: Book contributions that were available freely (as PDFs, e.g. on the author's website) become inaccessible to a large percentage of potential readers as a result of traditional publication. But for authors, association with a high-profile label is more important than accessibility of their book to readers, because they build their careers on the high-profile labels. In actual fact, the book is then often read in its pre-publication PDF version, which is still circulating underground. So the main purpose of the costly typesetting, printing, shipping and library storage of books is not to give readers access to its content (which they get otherwise), but to ensure the author's career. This situation is patently absurd, and it is very urgent to do something about it. This is what we want to do in this project.

2.1.1.2 Open-access publication

It has often been noted that open-access (OA) publication is more efficient for science, whether on a Green OA model, a Gold OA model (primary publication open-access), or a "Platinum" OA model (by which we mean Gold OA with no author processing charges (APCs), cf. Eve, 2012). The humanities differ from other fields in that APCs are not common. Prestigious publishers usually do not charge authors and requiring authors to pay would mean that we would lose many high quality submissions. Furthermore many books are written by researchers without access to external funding and these would be systematically excluded by author pays models.

The OA model, which is already widely used in journals in the sciences, can also be transferred to books in the humanities. We believe that the discipline of linguistics, which among the

humanities disciplines is the one that tends to behave the most like the sciences, is ideally suited to lead the way to open-access book publication.[1] In our experience, linguists are much less reluctant to accept the dissociation of high-quality content from paper format than practitioners of other humanities fields. But currently there are only very few open-access book series in our field, and the significance and output of these is quite marginal. No well-known linguist is involved in any of these:

- BELLS (Bergen Language and Linguistics Studies)[2]

- Aisberg[3]

- OApen[4]

- Asia-Pacific Linguistics (A-PL)[5]

College Publications has a linguistics series run by Ruth Kempson and Robin Cooper[6], but they published only two books, one in 2008 and another one in 2011. The volume from 2008 is a volume that was edited by the series editors and the one from 2011 is also an edited volume. The series editors seem not to be active in acquisition and the whole enterprise is hardly visible in the field. College Publication sells printed books and allow authors to put their book at their web page *once sales have achieved profitability.*[7] So it is Green OA with an embargo period.

It is the purpose of this project to set up a publication unit for Gold OA publications with the name *Language Science Press* that is run by academics (at the beginning by us).

We believe that it is possible to publish scientific books in our field at a much lower cost than traditional publishers, since we will distribute tasks among our community members (see Section 2.1.2.2 and 3.2.3), and that it is more sensible if books are published by academic institutions.

It is generally agreed that the transition from toll access to open access is the main problem. This is not only an economic, but also a social issue: Scholars are reluctant to switch to new modes of publication because no prestigious alternatives to the traditional publishers exist so far. It is here that we can provide a hopefully convincing alternative.

2.1.1.3 Costs of traditional and open-access publishing

Traditional publishers of printed scientific books have the following costs:

1. copyediting

2. typesetting

3. metadata creation

4. printing

5. shipping

6. advertising

---

[1] For example, a recent survey by the Societas Linguistica Europaea found that 93 % of the respondents preferred to access the society's journal electronically.
[2] https://boap.uib.no/index.php/bells
[3] http://aisberg.unibg.it/handle/10446/27154
[4] http://www.oapen.org/search?f1-subject=Humanities;f4-subject=Language::linguistics
[5] http://pacling.anu.edu.au/
[6] See http://www.collegepublications.co.uk/cml/.
[7] http://www.collegepublications.co.uk/about/. 24.09.2012.

  7. acquiring manuscripts and getting manuscript reviews

With e-books, (4) and (5) have become irrelevant, and in the internet age, (6) and (7) can be done more efficiently by the scholars themselves, who have the necessary contacts to their colleagues anyway. For example, we already have an agreement with the Linguist List that we can announce our books to the 25,000+ members of this mailing list. An example of a successful OA project in the journal sector is the journal FQS[8] which was founded by Dr. Katja Mruck (now Freie Universität, Berlin) and initially financed by the DFG. This project is now run completely by the community. The FQS has more than 17.000 registered readers (see also DFG press release, 2011). FQS is hosted by the Center for Digital Systems (CeDIS) at the FU, which will be our hoster too in case this project is financed.

Many publishers do not do (7) on their own but have established series for which the series editor and the editorial board does the work. The only contribution of the publisher is help in the organization of the workflow (software). Moreover, most publishers have not been providing copyediting services (1) for several decades, so this is something that is no longer felt to be necessary for highly specialized works.

So open-access e-books need (2)–(3) plus

  8. software for book publication, that is, the organization of reviewing, (1), (6) and (7)

  9. web presentation and permanent hosting

  10. link to a print-on-demand service

Since all research institutions have sufficient server space (9), and scientific libraries can easily provide the service of metadata creation (3), the main issues that we focus on in this application are:

  2. typesetting

  7. acquiring manuscripts and getting manuscript reviews

  8. software for book publication (organization of reviewing, (1), (6) and (7))

2.1.2 Our activities in preparation for this application

2.1.2.1 Community Organization

Book publishers get their reputation from the books that were previously published by them, but a new initiative has to secure the support of important figures in the field in order to attract high-quality contributions. In this we have been very successful.

In August 2012, we started the initiative "Open Access in Linguistics", by first building up a network of supporters including prominent linguists like Steven Pinker, Charles Fillmore, Adele Goldberg, Ivan Sag, Christiane Fellbaum, Wolfgang Klein, Stuart Shieber, and Steve Abney to name just some of them. The full list is available at our webpage.[9,10]

The stated goal was to set up a publication unit that publishes open-access books in (primarily theoretical) linguistics that have the same prestige as high-profile traditional publishers such as Oxford University Press. Every book will be published within a series, as is already common

---

[8] http://www.qualitative-research.net.

[9] http://hpsg.fu-berlin.de/OALI/sign/supporters

[10] Stefan Müller contacted 143 well-established linguists with one single mail. The return rate was over 78 %, with over 75 % of positive reactions. Given that high-profile researchers are very busy, this is a very good result and we expect a high percentage of those who did not reply to be sympathetic to the proposal as well. 67 of the prominent researchers agreed to work as reviewers.

practice in linguistics. The task of acquiring manuscripts and soliciting reviews for a series will be assigned to a team of series editors (one to three editors per series).

We put together an Advisory Board consisting of experts in the descriptive levels of Phonology, Morphology, Syntax, Semantics, Pragmatics and in the subdiciplines of Computational Linguistics, Typology, Historical Linguistics, Corpus Linguistics, Language Acquisition, Conversation Analysis. The board has representatives from diverse schools (Minimalism, Construction Grammar, Head-Driven Phrase Structure Grammar, Cognitive Linguistics). See Section 5.2 on collaboration. The Advisory Board helps make decisions regarding the establishment of new series. We plan to establish series in the main fields (descriptive levels) of linguistics and in the subdisciplines. The series editors will use the software that we provide (see Section 2.1.2.3), but apart from this work independently. Those who are responsible for a series also have to provide resources for proofreading and typesetting (but see Section 2.1.2.2).

2.1.2.2 Supporter Databases and Mailing Lists

Stefan Müller organized the initial list of 100 prominent supporters. His group implemented a data base for registration that allows supporters to register without human interaction. The database collects email adresses of potential authors and their book projects, it collects reviewers together with their area of expertise, volunteers that want to help in typesetting and/or software development, and it collects volunteers for proofreading.

44 out of 106 researchers who registered independently via the web form also registered as reviewers and hence can be contacted by series editors. 16 out of 106 researchers registered as typesetters, 38 as proof readers, and 43 as developers via the web form.

It is also possible to register as a reader. Registered readers will be informed about new books via newsletters.

2.1.2.3 Software, Web-Hosting and Network Infrastructure, and Workflow

We will use Open Monograph Press (OMP) by the Public Knowledge Project (PKP) as software for editorial work (reviewing, copy editing, proofing), for providing the catalog of our books and for sending out newsletters about new books to registered readers.

OMP has several roles that users can play: Press Manager, Series Editor, Press Assistant, Reviewer, and Author.[11] Press managers, series editors, and press assistants can assign reviewers and proof readers to submissions and thereby ensure that they can access only the workflow stage and monograph to which they have been assigned. The user manual describes the workflow as follows: Submitting a manuscript is a four-step process. Once the submission is complete, a press or series editor takes responsibility for the submission and initiates the submission's journey through the publication workflow: An Internal Review may be conducted, followed by an External Review; each of these optional stages involves its respective set of reviewers. Following the review stages the manuscript (if accepted) enters the Editorial stage, where copyeditors are assigned and a fair copy is generated and checked by the author. Following the Editorial stage, the submission enters the Production stage. There, the entirety of the submission's information – including submission placement, metadata, contributors, chapters, and so on – can be reviewed and finalized. Different book formats (digital, hardback, paperback/softback) can be configured, and proofs can be uploaded for auditing and final editorial approval. Once the proofs are approved and the catalog information checked and approved for publishing, the book is published online. Throughout the workflow, participants for each stage can be added and removed from the submission workflow by the series editors.

PKP is a well-established software developer who is also responsible for the Open Journal System (OJS), which is used by many OA journals, and the conference management system

---

[11] http://pkp.sfu.ca/wiki/index.php/OMP_Userguide. 23.02.2013.

Open Conference Systems (OCS). The software is open source and built and maintained by an international community which ensures maintenance even after the initial funding by the DFG. PKP and their local partners ensure the quality of code extensions that are added and modifications that are made to their systems. This regime of working ensures that all innovations and extensions are immediately available to other projects and disciplines as well.

We have a test installation of OMP[12] and first experiences with the workflow. The scenario was tested with one series, two categories, and two books. This installation will also be used for software development and further testing. The production system, that is, the software and the files that are used for the actual production of the books will be hosted by the Center for Digital Systems (CeDIS) at the Freie Universität Berlin. The CeDIS is an experienced partner of PKP and an experienced hoster of, service provider for, and developer of PKP software (see Section 2.1.1.3 on the success story of the OJS-based journal FQS).

The ZEDAT will provide the computing infrastructure. The FU is a large university with more than 33,000 students and the ZEDAT is a highly professional team that provides stable network connection, power supply and general infrastructure.

2.1.2.4 Long-Term Storage, Metadata, Maintenance

We set up a cooperation with the FU Library, which agreed to take care of cataloging the books that we produce and provides ISBNs and DOIs. The books will be stored on the document server of the FU, which is in the long term backup system and the safety of the data is guaranteed. In addition our library transfers copies of the files to the German National Library, which also is responsible for long term archiving.

2.1.2.5 LaTeX Setup

As we saw in Section 2.1.1.3, typesetting is the major cost involved in open-access book publication. But fortunately, many linguists already use LaTeX, a typesetting system that we will use to produce the final PDFs for the print on demand versions of our books.

Stefan Müller is an experienced LaTeX user. He typeset four published books and four book manuscripts. The two books that are crucial for the enterprise are the one on Persian (Müller et al., In Preparation) and the one on grammar theory (Müller, 2013). Persian is written in Arabic script, that is, right to left and has ligatures (characters that depend on the characters next to them), which is an enormous challenge for all latin-based computer environments. (1) provides an example for a glossed Persian sentence:

(1)   او مرد را دوست نخواهد داشت.
U   mard rā   dust   naxāhad dāšt.
He/she man   DOM friend NEG.want have
'He/she will not love the man.'

In another project Müller works on Mandarin Chinese. Again, the font issues are solved, as the example in (2) demonstrates.

(2)   狗   叫   了。
gou3 jiao4 le
dog   bark ASP/CRS
'The dog is barking/The dogs are barking.'

We are using XƎLaTeX, which allows us to use installed system fonts and hence avoids font problems.

---

[12] http://hpsg.fu-berlin.de/omp/index.php/LingPub/

The grammar theory text book discusses the major linguistic theories and uses the respective LaTeX packages.

So the main packages that are needed for producing linguistics books with LaTeX are availible. Stefan Müller has thaught LaTeX classes to linguists in the past (two day workshop at the University of Leipzig) and will update the teaching material[13] in the near future.

2.2 Verzeichnis projektspezifischer Publikationen

Müller, Stefan. 2006. LaTeXfür Linguisten. Handout of a tutorial given at the University of Leipzig. http://hpsg.fu-berlin.de/~stefan/Lehre/LaTeX/, 11.12.2013.

Müller, Stefan. 2012. A Personal Note on Open Access in Linguistics. *Journal of Language Modelling* 0(1), 9–39.

3 Ziele und Arbeitsprogramm

3.1 Ziele

As mentioned above, the goal of the project is to establish a publication unit (a kind of non-profit publisher) that is largely run by academic linguists themselves. Its publications will all be Gold open-access.

The process of the creation of this unit will be monitored by an economist who will help us to determine the exact costs of publishing and setup a sustainable business model.

Above, we mentioned the following two issues as the main challenges for publishing scientific books:

- quality control
- typesetting

These two points are related to the following tasks/needs that were identified in Section 2.1.1.3:

7. acquiring manuscripts and getting manuscript reviews
2. typesetting
8. software for book publication (organization of reviewing, (1), (6) and (7))

In order to address these issues we have the following goals:

1. organization of the community to ensure quality control (setup of the advisory board, establishment of series)
2. setup of layout and style files that is needed for book production and conversion into other formats
3. editorial work, conversion, and type setting
4. establish standards and develop software for conversion of LaTeX into XML for e-book production
5. extension of the OMP software (open reviewing and versioning, group features)
6. setup of OMP, customization of layout, training, maintenance and support
7. develop a business model

---

[13] http://hpsg.fu-berlin.de/~stefan/Lehre/LaTeX/.

3.2 Arbeitsprogramm

3.2.1 Organizing the Community/Quality Control

Our goal is to establish Language Science Press as a service that is run by the community and by research institutions.

We already begun to activate those members of our community that want to take part actively in an open access and open source project. We expect the first series to be founded in March 2013. The series editors are established linguists that guarantee prestige and will be responsible for building the brand. The editors of series that are proposed so far are experienced as editors and worked in large-scale projects. For instance Martin Haspelmath was responsible for the World Atlas of Language Structure[14], Stefan Müller edited the Proceedings of the annual conference on Head-Driven Phrase Structure Grammar (HPSG)[15] for the past ten years, Miriam Butt is a very active editor responsible for the proceedings of the annual conference on Lexical Functional Grammar (LFG)[16] and several volumes that were published with CSLI Publications, Stanford.

As was mentioned in Section 2.1.2.2, we already have a large pool of potential reviewers and this pool is extended by people who register via our subscription page (Section 2.1.2.2).

We will use OMP to build a database with active reviewers. We will evaluate reviewers with respect to the time that reviewers took to review manuscripts and the quality of their reviews. This will make it easier for editors to organize the reviewing process (see also Section 3.2.5.2).

In semester 0 (starting in March) the Advisory Board will be completed (Müller/Haspelmath) and the first series proposals will be evaluated. This work package is located before the expected funding by the DFG and no costs will arise.

3.2.2 Setup of layout and style files for typesetting

The books will be typeset in LaTeX. We are currently working with a designer on the general design of the covers and the layout of the books. We will provide LaTeX templates that ensure that complex typographical elements like trees, attribute value matrices, and OT tableaux are typeset correctly (see Müller, 2006 for an overview of these elements). This work package will be done as preparatory work without DFG financing in semester 0–1 by Stefan Müller and Timm Lichte (Universität Düsseldorf).

Furthermore we design a logo and develop a concept and a cover design for various book series, an open, extendable system. The design concept will be developed by Dipl. Designer Ulrike Harbordt in close interaction with Prof. Barbara Schmidt (Head of Design, Kahla Thüringen Porcelain GmbH and visiting Prof. at the UDK Berlin). This work will be carried out in the near future (semester 0–1). No financing via this grant application is needed.

Also in semester 0–1 the LaTeX tutorial (Müller, 2006) will be updated. It will be continuously adapted and extended during the complete runtime of the project. The LaTeX tutorial will be augmented by a tutorial on efficient publication including efficient use of LaTeX in a book project and fully automatic copying of files on department webservers. This tutorial will enable younger researchers to write their books in a format that is ready for publication in Language Science Press.

In order to be able to efficiently convert MicroSoft Word files into LaTeX, we have to develop a Word template and provide a detailed description of how to use it. This is necessary since many authors in the humanities do not use the formatting tools of Word properly, which makes automatic conversion difficult, if not impossible. This Word template will be developed by a

---

[14] http://wals.info/. 21.02.2013.
[15] http://cslipublications.stanford.edu/HPSG/
[16] http://cslipublications.stanford.edu/LFG/

person hired for Berlin (see next subsection). The same person will work on the Word to LaTeX conversion once accepted manuscripts are submitted for publication. We will use professional software for the conversion (Word to LaTeX Converter, http://www.wordtolatex.com).

No positions are requested for this work package.

3.2.3 Editorial work, conversion, and typesetting

Not all linguists are using LaTeX to prepare their documents and we therefore will also accept submissions in LibreOffice and MicroSoft Word. The Microsoft Word to LaTeX conversion usually is the most cost expensive part of book production. We will use commercial software to transfer Word files into LaTeX. After this initial, automatic conversion step additional manual labor is necessary. Elements on which the automatic conversion failed have to be typeset from scratch. Graphics that are of unacceptable quality, as for instance trees that are drawn with a graphics program in a non-uniform way, have to be redone. In principle this is cost intensive and time consuming, but as was mentioned in Section 2.1.2.2, we ask linguists to register as typesetters and software developers. We already have a pool of 16 typesetters and 43 developers.[17] Series editors will also be expected to contribute resources to the work that is involved in producing the respective books.

Linguistics is in a unique situation among the disciplines in the humanities since one branch of linguistics is computational linguistics, which is an intersection of computer science and linguistics. Researchers working in more formal frameworks like HPSG, LFG, and Tree Adjoining Grammar (TAG) use LaTeX to typeset their manuscripts anyway. So the expertise of these researchers can be used for the enterprise. Currently we have 16 volunteers that will help in the Word to LaTeX conversion.

Since we have 16 typesetters already and expect more to register, we can split the documents and let several type setters work on one document in parallel. This parallelization idea was discussed with experienced typesetters on the Tex-D-L mailing list and has been found tenable. With parallelization we expect to have much shorter turn around times than any traditional publisher, while ensuring better quality than those publishers that print from Word files directly. Of course there are no experiences with such a setting, so that a fall-back strategy is needed. We therefore ask for support for the initial phase of two years. The respective people (2 half positions) will help organizing the typesetting process and do typesetting themselves (see Section **??**).

3.2.4 Establishing standards and develop software for XML creation

In order to provide fully searchable documents that are annotated for content, we have to provide the documents in the XML format. This is also necessary for producing e-books in the ePub format. See also Section 3.2.7.3 on cooperation with the online platform CLEO, which can be used to generate an income for OA projects.

There are several projects that started to work on LaTeX to XML conversion as for instance LaTeXML[18] and tex4ht[19]. However, these have been written for other purposes. In order to use them, we have to standardize our LaTeX template (including heavily crossreferenced glossed example sentences as (1), trees, Attribute-Value-Matrices) and then implement a conversion that can be used as an extension of one of the existing LaTeX to XML converters.

In a separate work package we want to develop software that can be included into web pages to manipulate trees. Authors can then provide much richer example analyses without disturbing the clarity of the exposition. For instance an author can discuss the constituent structure of

---

[17] For current numbers see http://hpsg.fu-berlin.de/OALI/.
[18] http://dlmf.nist.gov/LaTeXML/
[19] http://tug.org/tex4ht/

an example using a figure like Figure 1. If the user wants to learn more, he or she may click a
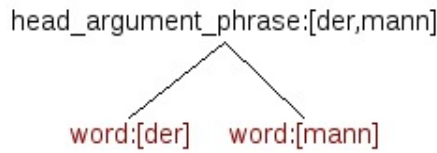

Figure 1: Simple tree showing constituent structure

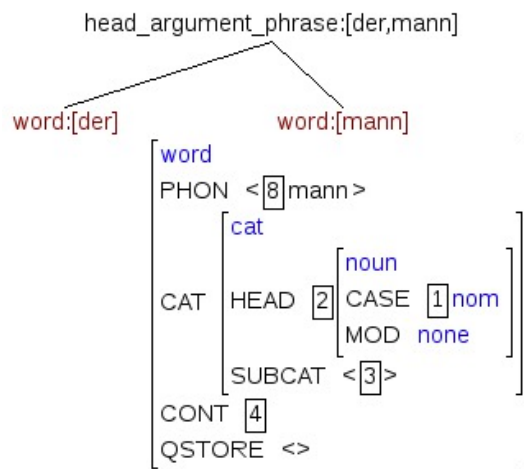node and display further information, as is demonstrated in Figure 2.

Figure 2: Expanded tree showing constituent structure and feature value pairs

Further extensions are the inclusion of audio files and movies (for instance for the documentation of sign languages).

Most of the tools and grammars that are developed by linguists are open source. The open source idea is deeply entrenched in this part of the community. We therefore expect enough interest in the Language Science Press project to be able to form a team of developers that is able to solve the tasks of the LaTeX to XML conversion and the software development that is described in the next section.

The basic work in LaTeX to XML conversion will be finished within a year. However, we expect new challenges to turn up when authors are using special styles for special layout elements in subareas of linguistics. For instance, when a new series starts.

In the second year we plan to work on extensions of the formats, that is, inclusion of animated trees and sound and video.

We ask for funding for a coordinator for this work package (two years, one PostDoc position).

3.2.5 OMP extensions

3.2.5.1 Open Reviewing and Versioning

It is often the case that comments of reviewers significantly improve books. Some reviewers invest a lot of time in reviewing. We will follow a suggestion by Pullum (1984, p. 266) and publish the names of the reviewers with the book, if reviewers want this (for incentives see

Section 3.2.5.2). This approach is also taken by the Edition Open Access[20] which is a book series run by the Max Planck Research Library for the History and Development of Knowledge. As Pullum pointed out in his column, standing in for a publication with one's reputation ensures that reviewers take reviewing seriously. But with the new ways of publishing we can take Pullum's ideas even further. We can build a web of trust. We can set up reviewing systems that keep the original submission of accepted manuscripts around and add the reviews. This can be done even if reviewers decide to stay anonymous. It is done successfully in other disciplines as for instance in Geoscience, Economy, and Biology (Pöschl and Koop, 2008) (see for instance the journal of Atmospheric Chemistry and Physics[21]).

In addition to the two reviewers that are determined by the series editors the whole community is free to comment on manuscripts once they are accepted for publication. The book will be available as soon as it is accepted but will be finalized after some discussion period. This will further improve the book.

The advantages of such an open review process and versioning are:

- full transparency

- credit to the earliest published version of ideas

- higher quality of submitted manuscripts (Pöschl and Koop, 2008, p. 106)

- higher quality of the final product due to better reviews and community feedback

- credit to the work of reviewers

3.2.5.2 Gamification

Open reviewing will be combined with aspects of gamification (Radoff, 2011). Websites that use gamification give users points and awards for their activities. An example is stackexchange.com. Stackexchange has a list of rated questions and answers and you get credits for asking and answering questions.[22] With such a system in place, series editors and members of the community can see who is an experienced user of this system. In addition to credits for activity there are certain badges[23] that are awarded to users for reaching certain goals. For instance users of stackexchange.com get a badge for editing 500 questions/answers. There are certain thresholds for user privileges that are assigned automatically by the system.

In our system points will be awarded for

- submitting manuscripts

- reviewing manuscripts (as a selected reviewer)

- reviewing manuscripts (as a community member)

- working in the advisory board

- working as a series editor

- proof reading

---

[20] http://www.edition-open-access.de/. 16.02.2013.

[21] http://www.atmos-chem-phys-discuss.net. 16.02.2013.

[22] See http://tex.stackexchange.com/questions/82693/url-field-inherited-despite-separate-cross-reference-item for an example.

[23] http://tex.stackexchange.com/badges. 16.02.2013.

While selected reviewers have to comment on the whole manuscript, community reviewers may comment on parts. The community may vote reviews both by selected and by community reviewers up.

Bonus points will be given for:

- submissions of manuscripts to the open review process, which will be optional in principle

- speed of reviewing

- transparency (anonymous publication of the review)

- transparency (non-anonymous publication of the review)

The gamification approach allows younger researchers to enter the game without being chosen as a reviewer by some series editor. PhD students often know their field very well, sometimes better than established researchers. An open reviewing system with community participation can activate additional expertise and add to the reputation of younger scientists.

Of course the reviewing work that resulted in the rejection of manuscripts has to be credited without revealing the identity of the reviewer. Anonymity is a problem in the beginning only. It can be guaranteed by introducing delays for the respective gratification.

In the case of manuscripts of bad quality, that is, manuscripts that would require a lot of work on the reviewer's side, it could be the case that nobody is willing to review the manuscript. This can be either accepted by the author as a rejection or the author could increase the motivation for reviewing by setting a 'bounty'[24]. Bounties can be set on systems like stackexchange to increase the priority of a question. Those who answer the question will get a bigger number of credit points and the person who asked the question has to 'pay' with some of his or her credit points (=reputation). If the manuscript has some good ideas in it, reviewers eventually will be willing to invest a lot of time in a manuscript. The extreme version of the 'bounty' idea is of course co-authorship.

We ask for a full position (PostDoc) for a coordinator of the two subpackages mentioned in this section (open reviewing/versioning and gemification).

3.2.6 Setup of OMP, customization of layout, training, maintenance and support

As was mentioned the OMP software will be hosted by the CeDIS. The CeDIS helps in setting up an OMP instance, adapt its layout and provides training to series editors and press managers.

For this workpackage which runs over two years we need 3/4 of a full position. We have to compete with market prices and it is difficult to find qualified staff. Therefore it is necessary to pay a PostDoc equivalent salary.

There is a second proposal in the program "Wissenschaftliche Monographien und monographische Serien im Open Access" by Prof. Martin Emmer, Freie Universität Berlin. Prof. Emmer also asks for 3/4 of a position. In case both proposals are accepted, we expect synergy effects and therefore would only need 1 full position in total for both projects (1/2 position für each project).

3.2.7 Business Model

3.2.7.1 Products and brand name

Books will be provided as PDFs and later also as e-books and HTML (see Section 3.2.4).

Our publications will be cited as follows:

---

[24] http://meta.stackoverflow.com/questions/16065/how-does-the-bounty-system-work. 16.02.2013.

Müller, Max. 2015. *A Grammar of Modern Sanskrit*. Empirically Oriented Theoretical Linguistics, No. 2145, Berlin: Language Science Press.

The books will be offered in our OMP catalogue. PDFs will be free and there will be links to Print On Demand services that offer good quality books for an affordable price.

3.2.7.2 Output

Our aim is to reach the following quantities of published books:

- 2014 20 books (5 series = 4 books per series)
- 2015 35 books (7 series = 5 books per series)
- 2016 54 books (9 series = 6 books per series)

This estimate is based on the fact that we have five series proposals by now with high profile series editors (see Section 5.2.2) and expect further series to be proposed soon. For instance one on methodological issues and one on neuro linguistics. We already have high quality submissions: three submissions to the typology series run by Martin Haspelmath and two submissions to the series run by Stefan Müller. We expect the number of submissions to rise considerably once the first books appear.

3.2.7.3 Running costs during the initial phase

Our publications will all be Gold open-access, and we hope that it will not be necessary to charge the authors ("Platinum open access"), even at later stages.

We set up donation pages and try to get larger sums of money from foundations and smaller amounts of money as donations from readers. We do not want to charge authors a processing charge, but those who have access to publication fonds or who have money available that was allocated for subsidizing printed books (Druckkostenzusschuss) are welcome to transfer an amount of money that is equivalent to author processing charges charged by commercial publishers (1500€–2500€). Of course we would not expect authors who submit to the series of Martin Haspelmath and Stefan Müller and who work in DFG-financed projects to donate, since in case this grant application is accepted the respective series will be financed by the DFG already. However, donations from authors who work in NSF or ANR funded projects are welcome and so are donations of authors who submit to series that are not supported by the DFG.

We already set up flattr, which is a micropayment system that allows users to donate small amounts of money on a monthly basis.

Additionally, we will examine the possibilities to cooperate with the Centre for Open Electronic Publishing (CLEO), which is financed with 7 Mio € by the French state until 2019. CLEO offers HTML, e-books in the ePub format, PDFs and printed books to readers and libraries. CLEOs main goal is to provide sustainable financial support for OA publications but the platform is not limited to OA publications. They also work on digitizing books from commercial publishers. These books from commercial publishers and OA publications are part of the bundles that are offered to libraries. CLEO has a so-called freemium model (Mounier, 2011), in which HTML is freely accessible to all readers, while PDFs, e-books and PoD books are sold. One part of this freemium model is a special service bundle that is offered to libraries (premium access to journals and book series, campus statistics of the use of the resources, export of records in MARC formats, alerts and subscriptions, a telephone hotline and personal training). 66 % of the income generated by the library contracts are returned to the publisher (Language Science Press in our case) and 50 % of the income that is generated by selling books (e-book, PDF, Print on Demand).

3.2.7.4 Running Costs after the Initial Phase

As was explained in Section 3.2.1, our goal is to establish Language Science Press as a service that is run by the community and by research institutions, which have to provide certain infrastructure like libraries anyway. After establishing procedures for proofreading and typesetting, the costs should be very low. In any case we will ask series editors to shoulder as much of the work as possible on their side, so that the work that has to be done at Language Science Press is minimal (for first agreements see Section 5.2.2).

However, there will be the need for at least one contact person for running OMP and general coordination. This person could be employed by any research institution of the world and we are optimistic that we will be able to find support for this single position after two years.

The goal of this first work package is to determine the real costs of our community-based approach and to explore possibilities for financing it in a sustainable way. We would like a PhD student (1/2 position) to work on the business model for two years.

3.2.8 Summary of the packages

Semesters 1–4 describe the workplan for the work that is hopefully financed by the DFG. The distribution of the work over the semesters is shown in the following table. The numbers refer to people and B stands for Berlin and L for Leipzig.

As was outlined in Section 2.1.2.2 we hope to be able to get a lot of community support for tasks like proofreading, but until the full establishment of all procedures and the existence of reliable proofreaders it is crucial to have one designated person that can act as contact partner for series editors and authors and help them dealing with OMP and the text formats. The other positions in the table correspond to the work packages that were described in Section 3.2.

The table contains an x, if we do not need funding for a certain semester, a 1 if we ask for a full position and 1/2 or 3/4 of we ask for a half or three quarters of a position, respectively.

	Package	0	1	2	3	4	person	section
1	Establishment of series	x					Mü/Ha	3.2.1
2	LaTeX template	x	x				Mü/Lichte	3.2.2
3	LaTeX tutorial / publishing tutorial	x	x	x	x	x	Müller	3.2.2
4	project coordination		1	1	1	1	1 B	
5	editorial work, type setting		1/2	1/2	1/2	1/2	2 B	3.2.3
6	editorial work, type setting		1/2	1/2	1/2	1/2	3 L	3.2.3
7	LaTeX/XML conversion		1	1			4 B	3.2.4
8	enrichment of HTML/XML				1	1	4 B	3.2.4
9	extension of OMP, versioning/open review		1				5 B	3.2.5.1
10	extension of OMP, gamification			1	1	1	5 B	3.2.5.2
11	OMP support		3/4	3/4	3/4	3/4	6 CeDIS	3.2.6
12	business model		1/2	1/2	1/2	1/2	7 B	3.2.7

3.3 Umgang mit den im Projekt erzielten Forschungsdaten

There will not be any research data in the classical sense, but the books that we produce will be cataloged by the university library and by the German State Library and hence the sustainability is guaranteed.

The software that we use is Open Source (except the professional Word to LaTeX converter) and the software that we produce will be Open Source too. It will be archived by common code sharing initiatives like sourceforge.net.

4 Beantragte Mittel

5 Voraussetzungen für die Durchführung des Vorhabens

5.1 Zusammensetzung der Arbeitsgruppe

Stefan Müller will work in the project (management, software development, editorial work). Furthermore his secretary Frau Viola Auermann (E9) can be of assistance. The working group will contribute a Hilfskraft position (41 h/month). Furthermore Dr. Roland Schäfer (an experienced computational linguist, TV-L 13) will help in software development.

Martin Haspelmath's working group will contribute a studentische Hilfskraft for 41 h/month. Apart from this Hans-Jörg Bibiko and Robert Forkel will contribute technical assistance. Both of them have experience with web-publications in science, e.g. the World Atlas of Language Structures Online, the Intercontinental Dictionary Series, and other projects in linguistics.

5.2 Zusammenarbeit mit anderen Einrichtungen oder Wissenschaftlerinnen und Wissenschaftlern

Our project is a truly international enterprise that involves researchers from all continents. This section lists those researchers with which we will cooperate intensely.

5.2.1 The Advisory Board

We already installed an advisory board that will help in establishing new series. It currently consists of

- Prof. Artemis Alexiadou (Universität Stuttgart)

- Dr. Jim Blevins (University of Cambridge)

- Prof. Geert Booij (Leiden University)

- Prof. Miriam Butt (University of Konstanz)

- Prof. Ewa Dąbrowska (Northumbria University, Newcastle)

- Prof. Arnulf Deppermann (Institut für Deutsche Sprache, Mannheim)

- Prof. Nomi Erteschik-Shir (Ben Gurion University of the Negev)

- Prof. Martine Grice (University of Cologne)

- Prof. Laura Kallmeyer (University of Düsseldorf)

- Prof. Mary Esther Kropp Dakubu (University of Ghana)

- Prof. Anke Lüdeling (Humboldt-Universität zu Berlin)

- Prof. Sam Mchombo (University of California, Berkeley)

- Prof. Detmar Meurers (Eberhard-Karls-Universität Tübingen)

- Prof. Rachel Nordlinger (University of Melbourne)

- Prof. Steven Pinker (Harvard University)

- Prof. Friedemann Pulvermüller (Freie Universität Berlin)

- Prof. Stuart Shieber (Harvard University)

- Prof. Dieter Stein (University Düsseldorf)

These researchers are leading scholars in their fields and some of them are known open-access activists. For instance Dieter Stein is responsible for e-language, the open-access journals of the Linguistics Society of America (LSA). Steven Pinker is the linguist with the second highest number of citations in google scholar in linguistics: 38739 as of 21.02.2013.[25]

5.2.2 Series Editors

We are starting with four series and more will be established during the next months.

- African Language Grammars and Dictionaries (Adams B. Bodomo)

- Empirically Oriented Theoretical Morphosyntax (Stefan Müller, Miriam Butt, University of Konstanz; Berthold Crysmann, CNRS, Paris; Laura Kallmeyer, University Düsseldorf)

- Studies in Diversity Linguistics (Martin Haspelmath, Max Planck Institute for Evolutionary Anthropology)

- Outstanding Dissertations in Phonetics and Phonology (Martine Grice, University of Cologne and Doris Mücke, University of Cologne and Taehong Cho, Hanyang University, Seoul)

The series editors already agreed to provide resources for proofreading.

5.2.3 Design/Developers/Typesetters

We will cooperate with Dipl. Designer Barbara Schmidt (Visiting Prof. at the Universität der Künste, Berlin and Head of Design Kahla Thüringen Porzellan), who will give advice in all questions concerning the graphics design of the book covers.

Timm Lichte, Heinrich-Heine Universität Düsseldorf will help to develop the LaTeX templates.

5.2.4 CLEO

As was mentioned in Section 3.2.7.3, we will cooperate with Centre for Open Electronic Publishing (CLEO) and produce XML that can be used in their system. We will closely cooperate with Pierre Mounier, who is the Associate Director of CLEO.

5.3 Apparative Ausstattung

The software and the files for the books (manuscripts, revisions, page proofs, final versions of the documents, …) will be hosted by the Freie Universität Berlin.

6 Erklärungen

7 Unterschriften

Prof. Dr. Stefan Müller   Prof. Dr. Martin Haspelmath
Berlin, den 27.02.2013   Leipzig, den 27.02.2013

---

[25] http://scholar.google.de/citations?view_op=search_authors&hl=en&mauthors=label:linguistics

Literatur

DFG. 2011. Pressemitteilung Nr. 53. http://www.dfg.de/service/presse/pressemitteilungen/2011/pressemitteilung_nr_53/, 27.02.2013.

Eve, Martin Paul. 2012. Open Access needs terminology to distinguish between funding models: Platinum OA/Gold NON-APC. https://www.martineve.com/2012/08/31/open-access-needs-terminology-to-distinguish-between-funding-models-platinum-oagold-non-apc/.

Haspelmath, Martin. 2012. The prestige of the publisher's brand name – an underestimated factor in science book publication costs. http://www.frank-m-richter.de/freescienceblog/2012/10/30/the-prestige-of-the-publishers-brand-name-an-underestimated-factor-in-science-book-publication-costs/.

Mounier, Pierre. 2011. Freemium as a Sustainable Economic Model for Open Access Electronic Publishing in Humanities and Social Sciences. *Information Services & Use* 31, 225–233.

Müller, Stefan. 2006. LaTeX für Linguisten. Handout of a tutorial given at the University of Leipzig. http://hpsg.fu-berlin.de/~stefan/Lehre/LaTeX/, 11.12.2013.

Müller, Stefan. 2012. A Personal Note on Open Access in Linguistics. *Journal of Language Modelling* 0(1), 9–39. http://hpsg.fu-berlin.de/~stefan/Pub/oa-jlm.html, 11.12.2013.

Müller, Stefan. 2013. *Grammatiktheorie.* Stauffenburg Einführungen, No. 20, Tübingen: Stauffenburg Verlag, second edition. http://hpsg.fu-berlin.de/~stefan/Pub/grammatiktheorie.html, 11.12.2013.

Müller, Stefan, Samvelian, Pollet and Bonami, Olivier. In Preparation. *Persian in Head-Driven Phrase Structure Grammar.* Empirically Oriented Theoretical Morphology and Syntax, Berlin: Language Science Press. http://hpsg.fu-berlin.de/~stefan/Pub/persian.html, 11.12.2013.

Pöschl, Ulrich and Koop, Thomas. 2008. Interactive Open Access Publishing and Collaborative Peer Review for Improved Scientific Communication and Quality Assurance. *Information Services & Use* 28(2), 105–107.

Pullum, Geoffrey K. 1984. Stalking the Perfect Journal. *Natural Language and Linguistic Theory* 2(2), 261–267.

Radoff, Jon. 2011. *Game On: Energize Your Business with Social Media Games.* Wiley.