FROM THE SAVANNAH TO THE CLOUD: BLOGGING EVOLUTIONARY LINGUISTICS RESEARCH

RICHARD LITTAUER

Computational Linguistics Department, Universität des Saarlandes
Littauer@coli.uni-saarland.de

SÉAN G. ROBERTS, JAMES WINTERS, RACHAEL BAILES

School of Philosophy, Psychology & Language Sciences, The University of Edinburgh
S.G.Roberts@sms.ed.ac.uk, J.R.Winters@sms.ed.ac.uk , R.L.Bailes-2@sms.ed.ac.uk

MICHAEL PLEYER

Department of English Literature and Linguistics, Ruprecht-Karls-Universität Heidelberg
Michael.Pleyer@hggs.uni-heidelberg.de

HANNAH LITTLE

Artificial Intelligence Laboratory, Vrije Universiteit Brussel
Hannah@ai.vub.ac.be

Over the last thirty years, evolutionary linguistics has grown as a data-driven, interdisciplinary field and received accelerated interest due to its adoption of modern research methodologies. This growth is dependant upon the methods used to both disseminate and foster discussion of research by the larger academic community. We argue that the internet is increasingly being used as an efficient means of finding and presenting research. The traditional journal format for disseminating knowledge was well-designed within the confines of print publication. With the tools afforded to us by technology and the internet, the evolutionary linguistics research community is able to compensate for the necessary shortcomings of the journal format. We evaluate examples of how research blogging has aided language scientists. We review the state of the field for online, real-time academic debate, by covering particular instances of post-publication review and their reaction. We conclude by considering how evolutionary linguistics as a field can potentially benefit from using the internet

1. Language Evolution Research: Then and Now

Darwin published *On the Origins of Species* in 1859, and it was this date, more than any other, that can be said to have been the beginning of evolutionary linguistics research. The 1,250 copies printed in the first edition of *The Origins* were all called for by the time the book hit the shelves, and many editions soon followed. The main venues for scientific research were then books, dissertations,

and the new format of scientific journals. Today, the monograph is seen mainly as a chance for a broader discussion, and dissertations are more a mandatory requirement for a doctorate than one in any number of large expositions published in a career. Current, cutting-edge research is published mainly in journal articles and conference proceedings. For language evolution, there are several journals which are seen as most relevant to language research: e.g. *Nature, Science, PNAS, ProcRoyalSocB, Human Biology*. In the past few decades, there have also been an increase in language-evolution focused conferences, the largest and most prestigious of which is unquestionably Evolang, started in 1996 in Edinburgh by Hurford, Studdert-Kennedy, and Knight (1996). Evolang serves as the best place to become aware of ongoing research, to draw lines in the sand regarding competing theories, and to disseminate one's own work.

However, conferences occur infrequently; Evolang occurs only biannually. Journals, with their notoriously slow review and publishing process, can delay research from being published for an equally long amount of time. Monographs or other research books can take even longer. This presents a problem; with many different teams and researchers working on similar research, it stands to reason that knowing the state of the field today, and not yesterday, is a must. This is increasingly difficult given the amount of research in the field and the slow publishing times. Furthermore, it is difficult to disseminate work in this subfield to the rest of academia, either because of publishing delays, or because conference proceedings are often not read outside of their particular subfield. This presents a problem for an interdisciplinary field such as evolutionary linguistics, which draws from various areas of research such as anthropology, neuroscience, linguistics, computer science, biology, and the evolutionary sciences, among others.

The most recent Evolang 9, which took place in March 2012 in Kyoto, differed from previous iterations of the conference in the amount of dialogue that took place online. This dialogue took place mainly on several web logs (*blogs*) set up by members of the conference, as well as on the micro-publication site Twitter.[a] This digitalisation of the conference is representative of a sea change in the field of post-publication review. Instead of waiting several months, or longer, for a response to occur in writing, reviews of papers and presentations were available within as short a time as a couple of hours later online, and in some cases, live messages were posted during a presentation itself. These responses included a post and discussion concerning linguistic replicators (Roberts, 2012, March 14), a refutation of a claim regarding the relevance of jellyfish eyes to complexity in evolution (Winter, 2012, March 21), and an online game that tested the working memory of chimps and humans (Roberts, 2012, March 22).

The ability to immediately review research, and disseminate it to a wide audience is novel, and has wide-ranging implications. In this paper, we discuss the

---

[a]Twitter. http://www.twitter.com

state of the field for online review and research dissemination using blogging, particularly involving the cultural evolution blog *Replicated Typo*.[b] We discuss particular instances of blogging, including some posts during Evolang 9, which show the possibility of immediate review. Finally, we envision how the interdisciplinary field of language evolution research may change with these developing technologies.

2. The State of the Field for Academic Linguistics Blogging

Blogging offers new opportunities for academics to collaborate with researchers from other fields and integrate data easily, as it is free from the funding, time, and field-specific issues of traditional publication. However, the power to publish results and theories freely, and to provide and receive rapid feedback has both positive and negative potential implications. On the positive side, new ideas can be presented and discussed easily with progress potentially much faster than a traditional journal peer-review, as was the case at the Evolang conference. The internet also provides a forum to engage the public about ongoing research, which facilitates dissemination beyond the halls of academic conferences and universities, and the paywalls of publishers and libraries. However, on the negative side, ideas and comments can appear in public and affect research without being properly assessed. While well-thought out responses can alleviate such concerns, immediate responses outside of the considerable time frame of normal publication, and outside of anonymous peer review, could dilute the impact of relevant research. Discussion on the internet can also suffer because some academics are tentative to post their thoughts online due to a lack of protection on intellectual property, and in the absence of a reputable, widely used or recognised way of referencing ideas presented on the internet.

Blogs are a useful source for discovering current research and a forum for open peer review, whether open (from the public) or closed (from co-authors on drafts). Linguistics blogs have been around for many years, such as Language Log (Liberman & Pullum, n.d.), but blogs dedicated to language evolution have emerged, too, such as Babel's Dawn, Shared Symbolic Storage, Culture Evolves!, Biolinguistics Blog, Replicated Typo (Bolles, n.d.; Pleyer, n.d.; Jordan, n.d.; Martn, n.d.; Winters, n.d.). While junior academics are prevalent in online discussions of linguistics, we note that well-established academics are also actively involved, e.g. Language Log, Culture Evolves!, Diversity Linguistics Comment, Vocalized/Vocalised, Language on the Move (Liberman & Pullum, n.d., Jordan, n.d., Haspelmath, n.d., Hall-Lew, n.d., Piller & Takahashi, n.d.).

ReplicatedTypo.com, a community blog centred around cultural evolution, founded initially by Masters students at the University of Edinburgh, has received over 225,000 hits in 3 years (counts gathered using Wordpress Analytics) and

---

[b]Replicated Typo. http://www.replicatedtypo.com

been awarded 8 editor's selections from ResearchBlogging.com.[c] As well as reporting on recent publications and conferences, basic introductions to Linguistics, evolution, mathematical modeling and animal signaling have been written. The interests of the multiple authors are varied, but the central research theme - evolutionary approaches to language and culture - remains the same. Their aims as science bloggers on Replicated Typo are: to highlight and discuss new research on language evolution; to engage with the general public by presenting language evolution research in an accessible way; to be a platform for open science research into language evolution. Discussions of posts on the blog have lead to revisions of research and discoveries of new avenues of research, as well as collaborations and clarifications of research by the authors of the studies reviewed. For example, a post about specific language impairment (Little, 2010, August 23) attracted comments from the original author, Dorothy Bishop (Bishop, 2010), and one post on the mapping of linguistic phylogenies to politics (Littauer, 2010, October 19) received a response from Simon Greenhill (Gray, Drummond, & Greenhill, 2009). Both Bishop's and Greenhill's responses allowed for public discussion between them and those commenting on their articles. This would either happen completely behind closed doors in the case of anonymous peer review, or over an extended period of time in the case of printed response letters.

3. Peer Review in the Blogosphere

There is no universal consensus on the method or acceptability of citing ideas from blogs. We argue that the devaluing of research and criticism appearing in open forums risks obstructing or hindering research. This is not merely a debate in Linguistics; examples are available elsewhere, such as the so called #*arsenicgate* scandal among evolutionary biologists (see Zimmer, 2011, December 2), where a paper chronicling arsenic-laced bacteria (Wolfe-Simon et al., 2010) was hyped by the media before publication, but condemned by many prominent science bloggers immediately afterwards for poor methodology. The authors responded by claiming that they would only respond to comments in peer-reviewed journals, sidelining the scientists who raised issues on public blogs (Shema et al, 2012).

A similar objection was raised by Jasmin & Casasanto (Liberman, 2012, March 17) in response to a LanguageLog critique by Mark Liberman (Liberman, 2012, March 13) of the robustness and size of their "Qwerty Effect" on language processing. The ensuing debate over statistical particulars included a serious concern over the ability of bloggers to damage the reputations of professional researchers, despite their peer-reviewed publications and regardless of the value of the critique. The quality standard of academic blogs is reasonable point of criticism, but could benefit from a consideration of the bigger picture. Arguably, this lack of enforced standards is a strength of the blog format, which simply results

---

[c]Research Blogging. http://www.researchblogging.com

in a higher overall volume of content, of varying quality. Blogs can therefore offer higher quantity and far better engagement, which is directly complementary to high-quality (though restricted) journal content. Though basic regulating standards are yet to be established, that blogs have enough wide-ranging influence to sway majority opinion against peer-reviewed articles is testament to their greater efficacy in disseminating knowledge and engaging the public. Misgivings about standards and plagiarism appear to concern blogs that are used to disseminating original works in progress, such as small-scale experiments and theoretical essays. While public access to experimental data, code or theoretical ideas may worry those protective of their contributions, public blogs are often a useful way to stake claim to an idea. Aside from this reassurance, it is important to highlight that methodological transparency is always desirable and more important for research itself than misguided sentiments over intellectual property; the idea that the research community need enact some "intellectual patent" is itself absurd and antithetical to academic advancement.

Another recent debate that has taken place online concerns Keith Chen's work on the influence language has on economic decisions (Chen, submitted), which was discussed in the workshop on constructive models (Roberts & Winters, 2012). Chen demonstrated a robust correlation between whether people speak a language with a morphologically marked future tense and increased levels of smoking, drinking and obesity and less saving and pension provision. The paper was put online, and sparked much interest and media coverage, one online journalist writing "Want to end the various global debt crises? Try abandoning English, Greek, and Italian in favour of German, Finnish, and Korean." (Fellman, 2012, January 1). Chen has recently also given a TED talk about this idea (see McManus, 2012, June 28). However, there have also been criticisms from bloggers. Geoffrey Pullum has criticised the typology that Chen uses (Pullum, 2012, February 9), and Mark Liberman has demonstrated with a simple model that correlations between unrelated cultural features are more frequent if the cultural features diffuse geographically (Liberman, 2012, February 12). Chen has also responded to these also through the medium of blogging (Chen, 2012, February 12). There have also been many comments from readers, with intellectual heavyweights are discussing an emerging hypothesis completely online before journal publication. Not only does this demonstrate the growing potential of blogs as a place where serious science is done, but also as an arena where academic debates can thrive with an immediacy that traditional journals do not offer.

4. Online Participation at EVOLANG

Ideas from Evolang were also discussed online. Andrew Smith (Smith, 2012) and Monica Tamariz (Tamariz, 2012) discussed the theory of cultural replicators. Both presentations were covered on Replicated Typo during the conference (Roberts, 2012, March 14, Roberts, 2012, March 20). This allowed people who were not

physically present at the conference to interact with the discussions. William Benzon, who was not present, wrote a lengthy response to the two views (Benzon, 2012, March 15). Together with technologies such as video conferencing, this changes the idea of conferences as insulated, static events to the possibility of conferences being platforms for a much wider range of interactions.

19 articles were published on Replicated Typo covering the proceedings of Evolang 9 (see Roberts, 2012, June 9), the majority of which were published before the conference had ended. Other blogs also covered the conference (e.g. Samuels, 2012, April 2, Alba, 2012, March 31). Twitter also recorded 128 tweets about the conference as it was running, many echoing the words of the speakers as they presented their research. This meant there were many more ways to interact with Evolang than ever before. One factor that has made this possible in the last few years is the proliferation of high speed internet and Wi-Fi as standard in conference venues and hotels. Tablet PCs, smaller laptops and smartphones also mean that it's easier access the web in the conference venue, and we noticed a shift in the apparent social acceptability of interacting with a computer while listening to a talk.

Continuous access to the internet also made it possible to check other research online during the talk and respond immediately with informed questions. In an extreme example, Gary Lupyan ran a short online experiment during the conference to counter a claim made by Massimo Piattelli-Palmarini (see Bodo Winter, this issue). In another example, an online experiment was coded during the conference that addressed Matsuzawa's demonstration of the working memory ability of chimps at Evolang (Matsuzawa, 2012), and the results are discussed in Quillinan & Roberts (this issue). Online connectivity is increasing the speed of interaction and academic progress at conferences. We also hope that this will encourage higher standards of academic accuracy in talks.

Experiments such as Quillinan & Roberts also allow for participation and engagement of the public in research that electronically crowd-sources participants. A similar style of public engagement and collaboration, known as citizen science, has been implemented in other fields to great effect in recent years. Projects such as Galaxy Zoo[d] and Whale FM[e] have gained popularity as the Research Excellence Framework in the UK calls for research to have greater engagement and impact.

Finally, Replicated Typo also offered authors the chance of posting short previews' of their talks a few weeks before the conference. This service can be beneficial because official abstracts for conferences are often not available before the conference starts, and the precise focus of a talk can change in the long gap between abstract submission deadlines and the first talk of the conference. In par-

---

[d]Galaxy Zoo. http://www.galaxyzoo.org/
[e]Whale.FM. http://whale.fm/

ticular, this gave less senior academics such as PhD and Masters students a chance to attract people to their talks, which was an important factor given the number of parallel sessions in this year's Evolang. In general, blogging allows researchers who are just starting to make a name for themselves and to network with other academics in the field. This is often a benefit for a young discipline like Language Evolution because researchers can often be physically isolated from others in the field. Indeed, Evolang 9 was the first time that some of the Replicated Typo bloggers had physically met, but the knowledge of each other from the blogosphere allowed them to immediately start interacting productively.

5. Academic Publication in the Future

Along with the rest of academia, the field of language evolution needs to change and respond to the growth of modern technology. Many journals publish their articles on the internet already with some journals existing solely online; for example, the Public Library of Science (PLOS) journals. Having journals online allows for the publication of code and data with the article, which journals like PLOS encourage, and this allows for other academics to replicate studies or build on existing work but online journals are still subject to the long process of peer review.

In recent years internet phenomena, such as blogs and social media, have started changing the way that some researchers are operating, especially as the traditional journal article is no longer the main method for disseminating research. As alluded to throughout this paper, there are many advantages to using the internet as a research tool. For legitimate change to occur, however, academics need to slowly embrace these new methods, by looking for research online, sharing their views in public prior to publication through commenting or social media, and blogging themselves. But while there are incentives, such as large amounts of readers, there are also drawbacks; "Unfortunately, most scientific output created on the Web goes unnoticed by current academic metrics, which measure scientific work published in 'conventional' academic literature." (Priem et al., 2010, October 26) The authors here believe that as academics use the internet as a tool for creation and dissemination of research, suitable new metrics will follow, such as Total Impact, which aggregates online presence from various sites.[f] With new metrics that matter to hiring boards (and thus, jobs), the cycle may feedback until online publication is the norm.

Blogs can facilitate real-time academic debate as we saw in the examples in the paper during Evolang. Blogs are also almost exclusively open-source and allow for interaction with all users of the internet which goes some way to engaging the public with ongoing research. As well as a increase of the utilisation of blogs, we foresee that in the future journals may attempt to subsume the blog-comment

---

[f]Total-impact. http://total-impact.org/

format of blogs and make it available to their subscribers. This however will have its disadvantages where blog post succeed as it doesn't engage the public, as it will keep discussion between academics, and also doesn't all for open access. On the other hand, quality control is well established in the journal system, and this will have to be better accounted for and dealt with within the blogosphere; how to plan for quality control here is an open question.

In summary, the field is changing, as can be seen by the active online engagement at the Evolang conference, and beyond on popular research blogs. We hope that researchers will embrace blogging as more than a side front to traditional publications and private discussions, and will see them as an opportunity to showcase their ideas, to engage the public, and to bring the field of language evolution further along.

Acknowledgements

References

Alba, C. (2012, March 31). Evolang 2012 at Kyoto. In *Sintaxi de Butxaca*. Retrieved Oct 12, 2012, from http://sintaxi.debutxaca.com/2012/03/evolang-2012-at-kyoto-i/.

Bishop, D.V.M. (2010). Overlaps Between Autism and Language Impairment: Phenomimicry or Shared Etiology? *Behavior Genetics, 40(5)*, 618-626. doi:10.1007/s10519-010-9381-x

Bolles, E. (Ed.). (n.d.). *Babel's dawn*. http://wwww.babelsdawn.com.

Benzon, W. (2012, March 15). In search of the wild replicator. In *Replicated Typo*. Retrieved August 24, 2012, from http://www.replicatedtypo.com/in-search-of-the-wild-replicator.

Chen, K. (2012, February 12). Cultural diffusion and the Whorfian hypothesis. In *Language Log*. Retrieved May 21, 2012, from http://languagelog.ldc.upenn.edu/nll/?p=3764.

Chen (submitted). The Effect of Language on Economic Behavior: Evidence from Savings Rates, Health Behaviors, and Retirement Assets. Submitted to *American Economic Review*. Retrieved August 24, 2012, from http://faculty.som.yale.edu/keithchen/papers/LanguageWorkingPaper.pdf.

Darwin, C. (1859). *On the Origin of Species by Means of Natural Selection, or the Preservation of Favoured Races in the Struggle for Life*. London: John Murray.

Fellman, B. (2012, January 1). Speaking and saving. In *Yale Alumni Magazine*. Retrieved Oct 13, 2012, from http://www.yalealumnimagazine.com/issues/2012_01/find_savings.html.

Gray, R. D., Drummond, A. J., & Greenhill, S. J. (2009). Language Phylogenies Reveal Expansion Pulses and Pauses in Pacific Settlement. *Science, 323(5913)*, 479-483. doi:10.1126/science.1166858

Hall-Lew, L. (n.d.). *Vocalized/Vocalised*. http://vocalised.wordpress.com.

Haspelmath, M. (n.d.). *Diversity Linguistics Comment*. http://dlc.hypotheses.org.

Hurford, J. R., Studdert-Kennedy, M. & Knight C. (Eds.) (1998). *Approaches to the Evolution of Language - Social and Cognitive Bases*. New York: Cambridge University Press.

Jasmin, K. & Casasanto, D. (2012). The QWERTY Effect: How typing shapes the meanings of words. *Psychonomic Bulletin and Review, 19(3)*, 499-504. doi:10.3758/s13423-012-0229-7

Jordan, F. (n.d.). *Culture Evolves!*. http://evolutionaryanthropology.wordpress.com.

Liberman, M. & Pullum, G. (Eds.). (n.d.). *Language Log*. http://languagelog.ldc.upenn.edu.

Liberman, M. (2012, February 12). Cultural diffusion and the Whoran hypothesis. In *Language Log*. Retrieved May 21, 2012, from http://languagelog.ldc.upenn.edu/nll/?p=3764.

Liberman, M. (2012, March 13). QWERTY: Failure to replicate. In *Language Log*. Retrieved Oct 12, 2012, from http://languagelog.ldc.upenn.edu/nll/?p=3837.

Liberman, M. (2012, March 17). Response to Jasmin's and Casasanto's response to me. In *Language Log*. Retrieved Oct 12, 2012, from http://languagelog.ldc.upenn.edu/nll/?p=3848.

Littauer, R. (2010, October 19). Mapping Linguistic Phylogeny to Politics. In *Replicated Typo*. Retrieved October 12, 2012, from http://replicatedtypo.com/mapping-linguistic-phylogeny-to-politics/2690.html.

Little, H. (2010, August 23). Specific Language Impairment, Autism, and Language Evolution. In *Replicated Typo*. Retrieved October 12, 2012, from http://replicatedtypo.com/specific-language-impairment-autism-and-language-evolution/1888.html.

Martn, T. (Ed.). (n.d.). *Biolinguistics blog*. http://biolingblog.blogspot.com.

Matsuzawa,T. (2012). Outgroup: The Study of Chimpanzees to Know the Human Mind. *The Evolution of Language: Proceedings of the 9th International Conference (EVOLANG9)*, 227233. WorldScientic. doi:10.1142/9789814401500_0030

McManus, E. (2012, June 28). Speaking of the future: Keith Chen at TEDGlobal 2012. In *Ted Blog*. Retrieved August 24, 2012,

from http://blog.ted.com/2012/06/28/speaking-of-the-future-keith-chen-at-tedglobal-2012/.

Piller, I. & Takahashi, K. (n.d.). *Language on the Move*. http://www.languageonthemove.com.

Pleyer, M. (n.d.). *Shared symbolic storage*. http://sharedsymbolicstorage.blogspot.com.

Priem, J., Taraborelli, D., Groth, P., & Neylon, C. (2010, October 26). *Alt-metrics: a manifesto*. Retrieved September 12, 2012, from http://altmetrics.org/manifesto/.

Pullum, G. K. (2012, February 9). Keith Chen, Whorfian economist. In *Language Log*. Retrieved October 13, 2012, from http://languagelog.ldc.upenn.edu/nll/?p=3756.

Roberts, S. G. (2012, March 14). Evolang coverage: Andrew Smith: Linguistic replicators are not observable, nor replicators. In *Replicated Typo*. Retrieved August 24, 2012, from http://www.replicatedtypo.com/evolang-coverage-andrew-smith-linguistic-replicators-are-not-observable-nor-replicators.

Roberts, S. G. (2012, March 20). Evolang coverage: More on linguistic replicators. In *Replicated Typo*. Retrieved August 24, 2012, from http://www.replicatedtypo.com/evolang-coverage-more-on-linguistic-replicators.

Roberts, S. G. (2012, March 22). The QHImp Qhallenge: Working Memory in Humans and Chimpanzees. In *Replicated Typo*. Retrieved October 13, 2012, from http://replicatedtypo.com/the-qhimp-qhallenge-working-memory-in-humans-and-chimpanzees/4947.html.

Roberts, S. G. (2012, June 9). Plenary talk videos from Evolang. In *Replicated Typo*. Retrieved August 24, 2012, from http://www.replicatedtypo.com/plenary-talk-videos-from-evolang.

Roberts, S. & Winters, J. (2012) Constructing Knowledge: Nomothetic approaches to language evolution. In L. McCrohon, T. Fujimura, K. Fujita, R. Martin, K. Okanoya, R. Suzuki, N. Yusa, *Five Approaches to Language Evolution: Proceedings of the Workshops of the 9th International Conference on the Evolution of Language*. Evolang9 Organizing Committee.

Samuels, B. (2012, April 2). Evolang 2012 Wrap-up. In *The Biolinguistics Blog*. Retrieved October 12, 2012, from http://biolingblog.blogspot.de/2012/04/evolang-2012-wrap-up.html.

Shema, H., Bar-Ilan, J., & Thelwall, M. (2012). Research Blogs and the Discussion of Scholarly Information. *PLoS ONE* 7(5): e35869.doi:10.1371/journal.pone.0035869

Winter, B. (2012, March 21). Evolang Coverage: Massimo Piatrelli-Palmarini's Plenary Talk. In *Replicated Typo*. Retrieved October

12, 2012, from http://www.replicatedtypo.com/evolang-coverage-massimo-piattelli-palmarinis-plenary-talk/4975.html.

Winters, J. (Ed.). (n.d.). *Replicated typo*. http://www.replicatedtypo.com.

Wolfe-Simon, F., Switzer Blum, J., Kulp, T., Gordon, G., Hoeft, S., et al. (2010). A bacterium that can grow by using arsenic instead of phosphorus. *Science* [online early access]. 10. Retrieved October 13, 2012, from http://www.sciencemag.org/content/332/60?34/1163.short.

Zimmer, C. (2011, December 2). Happy Birthday, #arseniclife. In *The Loom*. Retrieved March 14, 2012, from http://blogs.discovermagazine.com/loom/2011/12/02/happy-birthday-arseniclife/.