WALS in the university classroom: A review

KRISTINE A. HILDEBRANDT and OLIVER BOND

Abstract

The world atlas of language structures *(*WALS*) originally appealed to the linguistics community as a resource for research. However, the relevance of the feature chapters to teaching environments and the user-friendly nature of the Interactive Reference Tool also make it suitable for university classrooms. Based on our experiences using* WALS *in two typology courses at the University of Manchester and the School of Oriental and African Studies (SOAS), we provide a review of* WALS *as a teaching and learning tool, including both its successes and frustrations. We note some methodological and technical issues with using* WALS *in the university classroom, including problems of over- and under-sampling, and a lack of coverage on typological rarities. However, we have also found that* WALS *has much to offer instructors and students in terms of its breadth of topic coverage, the linkage of the feature chapters with course reading assignments, the wealth of genealogical, geographical, and bibliographic information on individual languages, and the hands-on experience that the Interactive Reference Tool offers students.*

Keywords: linguistic atlas, methodology, teaching

1. Introduction

In recent years, "Typology" as a taught undergraduate or (post-)graduate course has become more frequently attested at the university level. A recent survey via the listserv for the Linguistics Association of Great Britain (http://www.lagb. org.uk) revealed that at least five universities in the United Kingdom alone offer either undergraduate and/or post-graduate typology courses to their languages

and linguistics students on a regular basis.[1] In addition, some of these universities offer more than one taught course inspired by a typologically-informed teaching and research agenda, usually at different levels of student experience. This is exciting news for the sub-discipline, as it presents growing opportunities to introduce students to not only the so-called "traditional" areas of typological thought (e.g., language universals and core variable types), but also actively debated, cutting edge, and rapidly developing areas in contemporary typology, including issues on sampling, the introduction of non-traditional variables, and the role of (database) technology in typological research.

A recently released resource that adds to this growing opportunity in the typology classroom is *The world atlas of language structures* (*WALS*), edited by Martin Haspelmath, Matthew S. Dryer, David Gil, and Bernard Comrie (2005). As is well-known in the typology community, *WALS* is a relational database of crosslinguistic structural properties, crossing a broad spectrum of sub-disciplines within the field. It includes feature-chapter contributions from 55 authors (many are authors of two or more chapters), with data on 2,560 languages, a maximal sample size of 1,370 languages, and a core sample of approximately 200 languages. Originally released in 2005 as a book with a companion electronic Interactive Reference Tool on CD, *WALS* has recently been made more widely (and freely) available via the internet as a download of the original 2005 release or a more sophisticated web-based version: *WALS online* (Haspelmath et al. (eds.) 2008).

WALS originally appealed to the general linguistics community as a research resource (cf. Cysouw 2005; Cysouw & Albu 2005a, b on issues of using *WALS* for quantitative research; Cysouw et al. 2005 on interoperability issues inherent to the *WALS* database system). However, the type of information included in *WALS* and the largely user-friendly nature of the Interactive Reference Tool also make it suitable for use in research-based instructional domains.[2] This review is based on experiences using *WALS* as an instructional tool in two typology courses at the University of Manchester and the School of Oriental and African Studies (SOAS).[3] Each class was open to students at advanced undergraduate level and to students on taught Masters programmes. Since 2005 (2006 at SOAS) *WALS* has been integrated into these two courses

---

1. These are Newcastle University, School of Oriental and African Studies, University of St Andrews, University of Sussex, and University of Manchester.

2. For instance, *WALS* as a research tool was reviewed by Donohue (2006), who then noted that it would make a good teaching tool.

3. At the time of the composition of this review, the first author was employed as a lecturer at The University of Manchester (Linguistics and English Language) and the second author at SOAS.

in at least three dimensions: (i) as an instructor-oriented teaching resource, (ii) as a student-lead learning resource (via assessed exercises), (iii) as a student-organized tool for original research. Overall, we conclude that despite some methodological and technological problems, *WALS* has a great deal to offer instructors and students as both a teaching and learning resource. In the following two sections we outline the ways that *WALS* has been used in these dimensions, highlighting both its strengths and potential pitfalls. We first address the usefulness of *WALS* as a teaching aid in this context, and we then turn to *WALS* as a student learning and research aid.

2.   *WALS* as a teaching resource

2.1.   *Domain of use*

We define a teaching resource as one that aids the instructor in sourcing data and references for the purpose of constructing lectures and other course materials. Using *WALS* in this capacity potentially provides data in the form of examples from a variety of languages (from the chapter write-ups themselves), counts/frequencies gathered via an examination of single or combined feature-chapters, details of languages in which a particular linguistic phenomenon is attested (or noticeably absent), definitions from the chapter write-ups when appropriate, and references for primary sources. The map-making tool is also good for visual aid construction, especially when geography/areality is at issue. At Manchester, for example, a multi-week unit on the South Asian *Sprachbund* (including the testing of other potential areas) was enhanced by the mapping features available in *WALS*, including the ability to create tailored and overlaid maps with additional (or different) data points representing multiple areal traits. As such, *WALS* can be manipulated by the course convenor for a variety of means (i.e., to facilitate, enhance, and expand discussion of particular topics).

The breadth of data in *WALS* also makes it fairly adaptable to teaching a range of different topics. At SOAS, for example, its primary application has been to discuss phonological typology, based on the following motivations:

(i)   the chapters on the phonological characteristics of languages in the database are typified by large samples and are thus likely to have a more robust data distribution in *WALS*;

(ii)   employing the data in *WALS* (together with supplementary sources) is a good way of incorporating work on phonological typology into an introductory course, given that this is a burgeoning sub-field of typology that does not already feature heavily in introductory textbooks to typology;

(iii)   the localised distribution of certain (surface) phonological features provides an opportunity to introduce concepts like language contact early on in the course (e.g., the presence/absence of clicks in southern Africa, the

distribution of tone languages worldwide, etc.) to students who may not yet have a developed understanding of morphosyntax.

Given these opportunities, we now summarize some of the successes and frustrations we have experienced with *WALS* in this dimension.

2.2.   *Successes*

We found that in its most basic sense, *WALS* was useful for searches of languages that fit a particular criterion. For example, at Manchester, review units on implicational universals (e.g., the order of clause-level constituents with respect to ordering of elements in the noun phrase) were brought to life via (combined) surveys of the relevant feature-chapters (e.g., Dryer 2005d, e, f (Chapters 95–97)), as well as more detailed consideration of specific languages.

The Language Viewer in *WALS* is especially useful for quick access to genealogical/geographic information, in most cases complementing data from the *Ethnologue* (www.ethnologue.com). In particular, at times where internet access was restricted (e.g., lecture preparation while on an aeroplane), the interactive version of *WALS* made for a good portable resource for such data (however see below for genetic-affiliation clashes between *WALS* and the *Ethnologue*).

In general, we feel that the mapmaking device in *WALS* is sophisticated enough for general purpose-built maps (e.g., incorporation into handouts and/ or presentation software), and with a bit of patience, new data points can be added in individually tailored maps when necessary.

An important part of an advanced typology course is its ability to encourage students to engage in sophisticated and individually constructed research projects. To this end, we found the language profiles to be handy for references for recommended reading and data access (e.g., reference grammars and journal articles).

2.3.   *Frustrations*

Many of the problems we confronted with *WALS* as an instructional resource are similar to those noted by others in research environments (cf. Cysouw 2005 and Cysouw & Albu 2005b for discussion of issues like under/over-sampling, gaps, coding reliability, value overlaps, inaccurate value combinations, and feature dependencies). For instance, for instructional purposes, gaps in the sample prevent true frequency or ratio counts, and result in additional steps of having to search down the attestation and actual example of the criterion. For these reasons, *WALS* should be treated as a secondary data source (or as a location for inspiration and research-leads, and not as the ultimate source for all necessary data). As such, we usually found it to be a good companion or complement to a comprehensive grammar/dictionary library and more general typological

literature. We should also note along this line that there is a particular potential danger that students may misinterpret how reliable or comprehensive *WALS* is (given its MPI provenance) and, without guidance, it could become perceived by students as a replacement for carrying out more sophisticated typological work. *WALS* focuses on broad-scale typology, synthesising languages and constructions to a restricted set of types, and not fine-grained typology in which a large number of parameters are carefully considered. Therefore, it should be expressly understood by instructors (and university libraries) that *WALS* is NOT a substitute for regular and continued access to grammars, dictionaries, and other published typologically oriented material.

In terms of overall content, we found that intermediary features (i.e., typologically odd or rare phenomena) in *WALS* are sometimes eschewed for "bigger" features (or else simply lacking altogether). Now that *The Universals Archive* has begun to catalogue typological rarities (http://typo.uni-konstanz.de/rara/intro/), it would be useful to see a future version of *WALS* contain chapters on features like odd or multiple matching basic colour terms, rare speech sounds, odd (default) phonological domains, and "strange" derivational operations and morphology, for example. Such features would serve to expand students' appreciation of what can stand as a typological variable and how such variables may be understood in theoretical frameworks.

Although available in the main reference manual, details of the sampling methodology and practices of *WALS* or the principles underlying the genetic groups given in the "Genealogical Language List" (Dryer 2005g) are not available in the online or downloadable interactive versions of *WALS*. Given that the original *WALS* book is rather large and bulky, and almost certainly going to be kept in a reference-only (non-circulating) section of a library, this has made using this aspect of *WALS* impractical, particularly in institutions with large class sizes. Since awareness of the application of sampling practices is important to typology in terms of teaching and research, and given that *WALS* and *Ethnologue* have different genetic classification practices at times, some transparency concerning the different decisions in relation to this issue would be useful, and would also facilitate classroom discussion on issues and problems with genetic affiliation in typology.

3. *WALS* as a learning resource

3.1. *Domain of use*

For the purposes of this review, a learning resource is aimed at direct student accessibility and manipulation, where students can initiate a research project, learn about sampling practices and issues, master essentials of map-plotting and tailoring, and formulate or refine research questions and hypotheses. The use of a resource like *WALS* can promote enquiry-based learning, and encour-

ages students to think about different modes of discussing and displaying information.

At both Manchester and SOAS we have incorporated a mandatory introductory "familiarity" assessment involving *WALS*. In particular, we required students to:

(i)   use the search function for maps/chapters in order to count the number of chapters associated with a particular topic;

(ii)   use the Language Viewer to give the number of varieties listed for a genus, their names, and where they are spoken;

(iii) create a map predetermined by constraints set out by the lecturer;

(iv)   create a map contrasting two features of the student's choice: the skills they need to do this are covered in the previous activities; the main difference is that the student is required to think about what kind of features would be suitable ones to contrast or investigate.

This assessment familiarized the students with the "nuts and bolts" of the interactive component of *WALS*, and also introduced them to the companion chapters, the language reports, and the additional bibliographic resources that they themselves could tap into for their research projects. There are several benefits of using *WALS* in such a way. For instance, the small-scale, concise nature of each *WALS* chapter means that the authors have aimed at explicit and specific methodological decisions for the purposes of the respective (set of) research question(s); in the classroom, this has the knock-on effect of encouraging students to think carefully about the methodological decisions they make in their own work, forcing them to strive for focus and precision in formulating their own research questions. Similarly, the assessment allowed students to begin formulating their own reactions to the strengths and weaknesses of *WALS*, encouraging critical thinking in relation to both the software and the linguistic methodology underlying its design.

In Manchester, an additional *WALS*-based assessment complemented the reading and lecture-based unit on areality, where students were asked to investigate a (genetically diverse) region (e.g., Ethiopia, China, the languages comprising Mesoamerica or circling the Baltic Sea, etc.) for structural clues towards areality. In the spirit of evaluation of *WALS* as a research tool, the students were also asked to consider the limitations of using *WALS* towards the diagnosis of linguistic areas (elaborated on below). We now turn to a summary of the successes and frustrations in this area.

3.2. *Successes*

In a very general sense, the incorporation of *WALS* into a taught setting introduces a kind of "quasi-lab" activity to a course that is otherwise dominated by largely monologic lecture on relatively dense, highly specific, and at-times

dry topics. Both in Manchester and at SOAS, the *WALS*-centred activities took place out of the normal classroom and in a computer lab facility. This has had the benefit of allowing students to more actively engage with topics and each other, and to more openly ask questions and discuss issues that might otherwise overwhelm or intimidate them in a lecture-only environment. In the three years that this module has been structured this way, there has been a noticeable (and positive) change in the interactional dynamics between the instructor and the students once the first *WALS* assignment has been completed.

In another practical sense, students have also gained a basic familiarity in dealing with a simple database, and how to manipulate information contained within it to their own needs. In a way, this has provided the students with an instant sense of accomplishment in that they have been able to master a piece of technology in a relatively painless environment. In relation to other typology database projects, in Manchester in particular, the introduction of *WALS* as a database exemplar has paved the way for another unit on a critical investigation of the recent proliferation of typological databases (e.g., *The Surrey Morphology Group* projects http://www.surrey.ac.uk/LIS/SMG/group_projects.htm; *The Universals Archive* http://typo.uni-konstanz.de/archive/intro/; *XTone* http://xtone.linguistics.berkeley.edu/; *Stresstyp* http://stresstyp.leidenuniv.nl/; *Autotyp* http://www.uni-leipzig.de/~autotyp/; and the several other databases listed in the *Typological Database System Project* http://languagelink.let.uu.nl/tds/index.html, to name just some examples).

One obvious transferable skill that *WALS* offers students is an introduction to map-making, which students then have used in their typology essays to good effect. This dimension of *WALS* has been particularly useful in showing the geographical distances involved between lesser known languages where it is important they are not in contact with each other for the sake of the study. Interestingly we have found that some students who were not required to create maps as part of their assessment (e.g., in other modules) have still successfully used *WALS*-built maps in completely self-motivated circumstances. We found this to be an impressive expansion of a specific skill acquired in one domain into other research areas. As a disclaimer though, as may be expected, the quality of the outcomes and the use to which students put *WALS* has varied significantly according to individual student abilities. Given that the practical outcomes can at times be largely cosmetic (e.g., map-making), we note that the use of *WALS* for assessment purposes should not rely solely on this aspect of the software, but should rather be expanded into a consideration of the content as well.

3.3. *Frustrations*

At the start of each term, we have noticed a general sense of initial uncertainty of students when first being confronted with the *WALS* interactive resource. This is not necessarily a design flaw of *WALS per se*, as both the concepts of the course, as well as the technical specifications of this (or any) database can be intimidating at first, especially when independent (and critical) efforts are required. Indeed, the first piece of assessment is structured in a pseudo-scavenger hunt format in order to help familiarize students with the inner workings of *WALS* (interactive). We are hopeful that with the ongoing development of *WALS online*, including the potential addition of instructional resources, these concerns can be minimized even more.

We have also noticed that the teaching and research gaps noted in the above section also quickly become apparent to students through their own use of *WALS*. For example, students often comment in footnotes to the completed assignments and research projects that there were limitations to the application of *WALS* that hindered their map-making, or else limited the crosslinguistic comparisons that they wanted to include in a survey or case study. For example, they are interested in manipulating the feature criteria or else adding additional parameters to a feature-chapter. Related to this, the feature manipulation and map tool limitations quickly become apparent to students (e.g., they are not able to manipulate the presentation of geographic labels). In a sense, we see their frustrations on this level as both a benefit as well as a drawback to *WALS*, as such reactions indicate that they are thinking outside of the constraints imposed by this single piece of software.

Despite the thousands of datapoints in *WALS*, the internal organisation of these feature-chapters can at times be difficult to reconcile with data and predictions encountered in course readings. On the one hand, *WALS* nicely complements some discussions in the readings, for example Croft's exemplification of implicational universals regarding the ordering of adpositions and the noun with respect to genitives and the noun (2003: 58–59, cited from Greenberg 1966: 78, cf. Dryer 1992). He notes that the prediction of Prep & NGen and Post & GenN to the exclusion of the other two combinations is not exceptionless, and this overwhelming preference (with exceptions) is indeed encountered in a modified version of Chapters 85 and 86 (Dryer 2005a, b), shown in Map 1.[4] It is useful when queries done in *WALS*, such as this one, (at least largely) match up with main points of discussion in course literature.

---

4. This is also a nice way to illustrate the feature combination tool in *WALS*. Additionally, *WALS* even allows for an initial foray into Croft's (2003: 59) consideration of these exceptions by a language-specific examination of the exceptional cases in light of the distribution of adjective-noun orderings in Chapter 87 (Dryer 2005c).
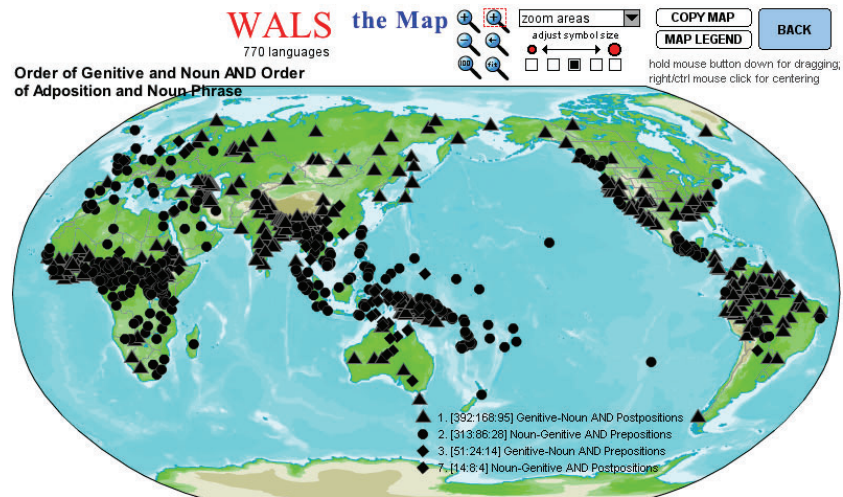
Map 1. *The order of the Genitive & Noun and Adposition & Noun Phrase*

However on the other hand, predictions encountered in other readings do not always synchronize so neatly with the features (and their codings) offered in *WALS*. One example of this is a unit on crosslinguistic diversity and stability, based largely on Nichols 2003. Her survey of structural diagnostics includes word order and the presence/absence of ergativity (easy to test in *WALS*), but she also includes the inheritance of basic vocabulary and the borrowing (or not) of personal pronouns. True enough, *WALS* contains some chapters on lexical patterns, but these are limited to a couple of body parts, colour terms, or else numeral bases. And while there are several chapters dealing with personal pronouns, none of these is assembled to allow a close examination of the actual morphological forms of the roots.

A similar frustration is there for students in a unit on areality. Upon a reading and discussion of Masica 2001, students quickly become aware of the problems associated with actually testing most of the *WALS* features for areal distributions. For example, the presence/absence of retroflex consonants, a strong diagnostic of areality in Masica's account, is not included in *WALS*. Another example is the presence/absence of (and reasons behind) direct object marking, which is less developed in languages of the Indian sub-continent, but more developed in northern India, Pakistan, and throughout the western Himalaya. This diagnostic is only indirectly observable in certain chapters of *WALS*. Once again though, we see such non-synchronizations as fuel for critical discussions

about research methodology, and we feel that the students still can benefit from these issues.

4.   Other comments and conclusion

A brief note on the book-plus-CD ROM vs. online version of *WALS* is in order here. Until recently, *WALS* was only available as a single unit to be purchased (at some considerable cost). In our cases, the university libraries or departmental resource rooms were the purchasers and holders of these materials, which made for limited access options and lowered the likelihood that students would take much advantage of *WALS* outside of structured classroom activities or university opening times. These constraints also precluded it being used in distance learning situations. The more recent availability of an online version is extremely beneficial, as students and instructors can download the Interactive Reference Tool to use on their personal computers or access the more intuitive web-version online. We also note that the newer online version of the database has some improvements over the CD ROM-only version, including the ability to build maps directly, access to the feature-chapters, more extensive instructions, useful software downloads, links to related projects, and a Newsblog. Again, this is great news for education purposes, and, in addition to these improvements, we would also like to suggest that links to teaching resources using *WALS*, as well as a link to manuscripts based on *WALS* data also be made available on the *WALS online* site.

Beyond the typology classroom, we wish to note in closing that one overarching strength of *WALS* is that it allows students at all levels of experience to appreciate that the study of crosslinguistic variation is not a necessarily frightening or impossible endeavour. Quite to the contrary, we have observed that once students are exposed to *WALS*, they frequently move on to (more) readily embrace case studies of features and languages from outside of their traditional areas of classroom training. Both the qualities and the problems of this programme thus provide equally useful methodological and topical lessons and opportunities, and as such, we see *WALS* as an incredibly valuable resource in any typology or crosslinguistic structure-survey course for students of linguistics.

Correspondence addresses: (Hildebrandt) Department of English Language and Literature, Southern Illinois University Edwardsville, 2227 Peck Hall, Edwardsville IL 62026, U.S.A.; e-mail: khildeb@siue.edu; (Bond) Department of Linguistics, School of Oriental and African Studies, University of London, Thornhaugh Street, Russell Square, London WC1H 0XG, United Kingdom; e-mail: ob1@soas.ac.uk

References

Croft, William. 2003. *Typology and universals*. 2nd edn. Cambridge: Cambridge University Press.

Cysouw, Michael. 2005. Using WALS: Prospects of quantitative approaches to linguistic typology. Paper presented at the Seminar on the World Atlas of Language Structures, Max-Planck-Institut für evolutionäre Anthropologie, Leipzig. Slides available at http://email.eva.mpg.de/~cysouw/pdf/cysouwWALSINSTITUTE.pdf (17 June 2008)

Cysouw, Michael & Mihai Albu. 2005a. Preparing WALS for quantitative analyses. Paper presented at the Workshop on the World Atlas of Language Structures, Padang, Indonesia. Slides available at http://email.eva.mpg.de/~cysouw/pdf/cysouwWALSQUANT.pdf (17 June 2008)

Cysouw, Michael & Mihai Albu. 2005b. Revamping the World Atlas of Language Structure. Discussion meeting on typological databases, Max-Planck-Institut für evolutionäre Anthropologie, Leipzig. Slides available at http://email.eva.mpg.de/~cysouw/pdf/cysouwDATABASES.pdf (17 June 2008)

Cysouw, Michael, Jeff Good, Mihai Albu & Hans-Jörg Bibiko. 2005. Can GOLD "cope" with WALS? Retrofitting an ontology onto the World Atlas of Language Structures. *Proceedings of E-MELD 2005 Workshop on Digital Language Documentation.* http://emeld.org/workshop/2005/proceeding.html

Donohue, Mark. 2006. Review of Haspelmath et al. (eds.) 2005. *The Linguist List* 17. 1055. http://linguistlist.org/issues/17/17-1055.html

Dryer, Matthew S. 1992. The Greenbergian word order correlations. *Language* 68. 81–138.

Dryer, Matthew S. 2005a. Order of adposition and noun phrase. In Haspelmath et al. (eds.) 2005, 346–349.

Dryer, Matthew S. 2005b. Order of genitive and noun. In Haspelmath et al. (eds.) 2005, 350–353.

Dryer, Matthew S. 2005c. Order of adjective and noun. In Haspelmath et al. (eds.) 2005, 354–357.

Dryer, Matthew S. 2005d. Relationship between order of object and verb and order of adposition and noun phrase. In Haspelmath et al. (eds.) 2005, 386–389.

Dryer, Matthew S. 2005e. Relationship between order of object and verb and order of relative clause and noun. In Haspelmath et al. (eds.) 2005, 390–393.

Dryer, Matthew S. 2005f. Relationship between order of object and verb and order of adjective and noun. In Haspelmath et al. (eds.) 2005, 394–397.

Dryer, Matthew S. 2005g. Genealogical language list. In Haspelmath et al. (eds.) 2005, 584–644.

Greenberg, Joseph H. 1966. Some universals of grammar with particular reference to the order of meaningful elements. In Joseph H. Greenberg (ed.), *Universals of grammar* (2nd edn.), 73–113. Cambridge, MA: MIT Press.

Haspelmath, Martin, Matthew S. Dryer, David Gil & Bernard Comrie (eds.). 2005. *The world atlas of language structures*. Oxford: Oxford University Press.

Haspelmath, Martin, Matthew S. Dryer, David Gil & Bernard Comrie (eds.). 2008. *The world atlas of language structures online*. München: Max Planck Digital Library. http://wals.info (17 June 2008)

Masica, Colin P. 2001. The definition and significance of linguistic areas: Methods, pitfalls and possibilities (with special reference to the validity of South Asia as a linguistic area). *The Yearbook of South Asian Languages and Linguistics* 4. 205–267.

Nichols, Johanna. 2003. Diversity and stability in language. In Brian D. Joseph & Richard D. Janda (eds.), *The handbook of historical linguistics*, 283–310. Malden, MA: Blackwell.