

Linguistic patterns of academic Web use in Western Europe

MIKE THELWALL,¹ RONG TANG,² LIZ PRICE¹

¹*School of Computing and Information Technology, University of Wolverhampton (UK)*

²*School of Information Science and Policy, University at Albany, SUNY, Albany (USA)*

A survey of linguistic dimensions of Web site hosting and interlinking of the universities of sixteen European countries is described. The results show that English is the dominant language both for linking pages and for all pages. In a typical country approximately half the pages were in English and half in one or more national languages. Normalised interlinking patterns showed three trends: 1) international interlinking throughout Europe in English, and additionally in Swedish in Scandinavia; 2) linking between countries sharing a common language, and 3) countries extensively hosting international links in their own major languages. This provides evidence for the multilingual character of academic use of the Web in Western Europe, at least outside the UK and Eire. Evidence was found that Greece was significantly linguistically isolated from the rest of the EU but that outsiders Norway and Switzerland were not.

Introduction

In an era of increasing integration in the European Union and the resulting centralised funding for multinational scientific efforts, language has become an important issue for European scientific research (*Europa*, 2001, 2002a). With most top journals being exclusively English (e.g. *Garfield*, 1967; *Moed*, 2002), language is potentially a barrier to career advancement for the non-English speaking researchers. As evidenced in *Glänzel & Schubert's* (2001) study, shared language is one of the predictors of higher international research collaboration. Through an investigation into the collaboration between the top five producers of science, *Zitt et al.* (2000) found linguistic factors a potential obstacle to “the process of Europeanization in science” (p. 627). In fact English for Academic Purposes is a recognised research and teaching area (e.g. *Flowerdew & Peacock*, 2001). In Europe, there is a growing linguistic culture in education of multilingualism with English, so that English is embedded in the education process along with the national language(s) (*Hoffmann*, 2000). In a wider context, English has become such a widely used language for international communication that non-native outnumber native English speakers (*Crystal*, 1997).

Received November 12, 2002.

Address for correspondence:

MIKE THELWALL

School of Computing and Information Technology, University of Wolverhampton

Wulfruna Street, Wolverhampton WV4 4ST, UK

E-mail: m.thelwall@wlv.ac.uk

0138–9130/2003/US \$ 20.00

Copyright © 2003 Akadémiai Kiadó, Budapest

All rights reserved

There are countervailing tendencies, however, to the actual and potential 'linguistic genocide' of the smaller languages by the larger, initiatives to actively promote the use of less widely spoken languages (*Pennycook, 1996; Skutnab-Kangas, 2000*). The complexity of linguistic borders is well illustrated by the case of moving boundaries inside bilingual Belgium (*Treffers-Daller, 2002; Willemyns, 2002*).

Given the need to collaborate within the EU, an important question, then, is the extent to which scientists are managing to overcome the language barriers in their scholarly communications. UK and Eire researchers tend to grow up with the international language of science (although Welsh and Gaelic languages are also indigenous to these isles), but Greece has the burden of a different set of alphabets to master to operate effectively within the EU. Most such information exchange is by its nature transient and almost impossible to capture on a large scale. One important exception is the Web. Although probably not central to team building or day to day collaboration, university Web sites do host relatively informal scholarly information (*Wilkinson et al., 2003*) and represent scholarly output in an arena subject to considerably less coercion than formal publications (e.g. *Jiménez-Contreras et al., 2002*). Web links are particularly interesting as indicators of recognition by the author of one page to that of another, especially when between universities in different nations. Many commentators have predicted their future value as a new information source (*Almind & Ingwersen, 1997; Davenport & Cronin, 2000; Cronin, 2001; Borgman & Furner, 2002*). Given the availability of multilingual search options in Web retrieval tools, such as AltaVista, an opportunity is now presented to examine linguistic dimensions of academic Web use.

Related research

The two most salient results about Web links from the perspective of this paper are: counts of links between universities in the same country have been shown to associate significantly with measures of research productivity in the UK, Australia and Taiwan (*Smith & Thelwall, 2002; Thelwall & Tang, 2002; Thelwall, 2002a*); and that approximately 90% of national between university links are created for reasons associated with scholarly activity (*Wilkinson et al., 2003*). Both of these findings require extrapolation to be applied to sets of EU university links, but they set the context for this kind of Web link research. It must be admitted, however, that this research area is relatively young and much more must be found out about motivations for creating Web links if it is to have a significant future.

Other Web link research has investigated different methodologies for counting hyperlinks (Thelwall, 2002a) finding that counts based upon domains are more robust than those based upon pages, although when search engines are used to supply the raw data the page model is normally the only one available. It is also known that counts of links to journal Web sites can correlate with the citation based impact of the associated journal (Vaughan & Huysen, 2002) and that link counts vary by discipline (Vaughan & Thelwall, 2002). This underlines the potential for using hyperlinks to study scholarly communication. Web links have also been used from a communication networks perspective to study groups of Web sites (Park et al., 2002; Garrido & Halavais, 2002). The techniques applied have been able to analyse networks to extract concepts such as node centrality and the existence of coherent subgroups – “Cliques”. Web links are also extensively used from an information retrieval perspective (Brin & Page, 1998; Kleinberg, 1999), which underlines the fact that they can be mined for genuinely useful information.

Previous studies on EU universities Web sites (Thelwall et al., 2002) found that sizes of Web sites varied enormously, with many Eastern European university sites being significantly smaller compared to their western counterparts. Polanco et al. (2001) used techniques to cluster EU university Web sites, with raw data based upon co-linking. The clustered universities were on an individual rather than national level.

Linguistic factors in Web linking have been analysed once before (Thelwall & Tang, 2002) as part of a study comparing Mainland China and Taiwan. Although English was a widely used language in both national university Webs, accounting for 37% and 29% of pages respectively, there was no evidence that it was the language of choice for international link pages. Language issues across communities in the EU are of more general interest than just for academic communication, and those issues have been extensively studied in linguistics (e.g. Rash, 2002; Treffers-Daller, 2002; Willemyns, 2002). Moreover, on the Internet it was discovered that special linguistic variations of English are emerging (Crystal, 2001), although these are still essentially English, albeit with enhancements.

Research questions

This study addresses two specific research questions. Firstly, is there evidence that English is the standard language in the EU for the relatively informal melange of scholarly communication represented by Web links? Secondly, is there evidence that Web use significantly crosses linguistic boundaries within Western Europe? If this were not to be the case then the linking patterns are expected to be dominated by links

between countries sharing a common language, e.g. France – Belgium, Netherlands – Belgium, Germany – Austria – Switzerland. If other significant trends were to be found then this will be taken to be evidence of the importance of non-linguistic factors.

Methodology

Data collection

Advanced queries in the search engine AltaVista were used to obtain linguistic data on Web site interlinking. Search engines are problematic as a data source partly due to their reported unreliability (*Bar-Ilan, 1999; Rousseau, 1999*) although AltaVista appears to have become more stable recently (*Thelwall, 2001a*). Search engines can only report on pages that they indexed however, and this means that they will miss large areas of the Web (*Lawrence & Giles, 1999; Thelwall, 2002b*). These include dynamically generated pages in many cases, banned pages and others that are not linked to or directly reported to the AltaVista site (*Thelwall, 2001b*). Nevertheless, AltaVista's coverage of the academic Web appears to be much better than that of the commercial Web (cf. *Lawrence & Giles, 1999; Thelwall, 2001b*), and has been used to produce significant results for the national systems of university Web sites (*Thelwall, 2001b, 2002a*). We decided to use AltaVista in the context of this study, although readers are forewarned that the results should be treated with caution.

We restricted the study to the larger EU countries, since the character of linking to smaller nations appears to be different to that of large nations (*Thelwall & Smith, 2002*), in other words link counts do not scale linearly. To this set we added Switzerland and Norway, which are geographically contiguous and are hence useful for comparison purposes. For each one of these countries a list of its university Web site domain names was obtained, making 450 in all. These were organised into groups by nation. A program was then used to convert these into a series of 1024 Boolean queries for AltaVista, where the form of the queries was as follows.

(a) *Count the number of pages hosted by set i of country A that contain at least one link to any page in set j of universities in country B and language X.*

The maximum Boolean query size limitation of AltaVista made it necessary to use this complex type of query, with the larger nations' universities split into groups. A simplified artificial example is given below.

- Country A is Spain: set i is {ualv.es, ubc.es}
- Country B is Portugal: set j is {ulis.pt, unat.pt}
- Language X is Swedish

The query would then be as follows:

```
(host:ualv.es OR host:ubc.es) AND (link:ulis.pt OR link:unat.pt)
- select language = Swedish from drop-down menu
```

To find the total number of pages linking from Spain to Portugal in Swedish, the results of the queries above would be totalled for all different sets of Spanish and Portuguese universities.

The second type of query was to obtain the number of pages in a given country in a given language. These figures are needed for basic statistics and normalisation purposes.

(b) *Count the number of pages hosted by set i of country A in language X.*

Using the above country A and Swedish again the query would be as follows.

```
(host:ualv.es OR host:ubc.es)
- select language = Swedish from drop-down menu
```

All queries were submitted to AltaVista on the second of July 2002.

Data analysis

The techniques used to analyse the data are purely descriptive: network diagrams, as used in *Thelwall (2001c)* and *Thelwall and Smith (2002)*; and standard graphs. The network diagrams will be described in more detail because they are relatively new and include a novel variation of the previously used means of calculating the arrow widths.

The numbers behind each graph are to be the numbers of pages in universities in the source country that are in the specified language and contain at least one link to any page in a university in the target country. The objective of the network diagrams is to illustrate the relative use of each language for hosting international Web links to recognised universities in the chosen set of countries. As a result of this, the figures are normalised for the size of the source and target university Web site sizes by being divided by the total source page count and total target page count, both figures being the total count of all pages irrespective of language. The width of the arrows, then, represent the per page relative tendency to link between universities in the countries in the given language. The denominator of this calculation is necessary to stop the larger countries from swamping those of the smaller countries, which would obscure the linguistic trend.

Results and preliminary discussion

Linguistic variations amongst countries

Figure 1 shows clearly the dominance of English in the European academic context, accounting for 56% of all pages. The low showing of French is surprising, given its status as a major international language.

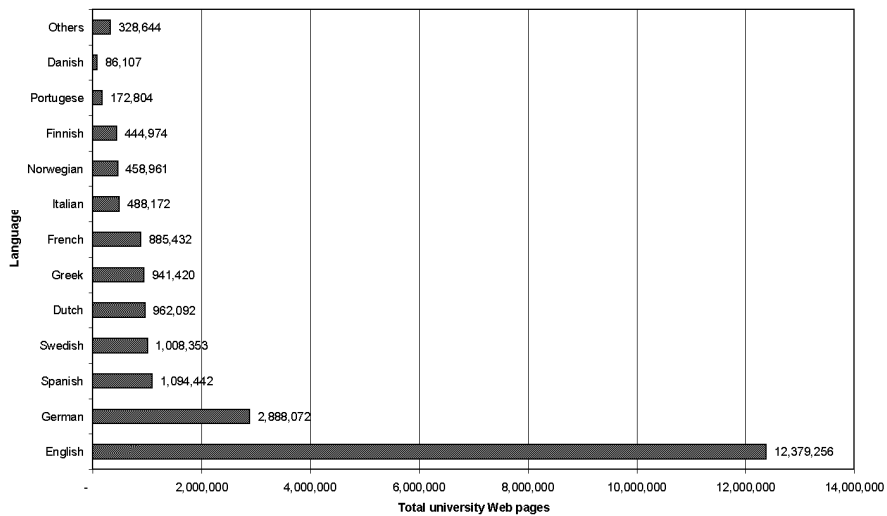


Figure 1. A breakdown of total pages by language, for all 16 EU countries covered

From Figure 2 it can be seen that bilingualism is the norm, with the exception of UK and Eire, where English is almost exclusively used. Greece is the most monolingual other than these two, with 85% of pages in Greek. Swedish is also a multinational language, with a significant presence in neighbouring Finland as well as a small presence in Denmark, Norway, Switzerland and Austria. Note that Belgium and Switzerland are both countries with three major languages extensively used. On the other hand, Italian does not feature in Switzerland despite its official status, perhaps due to national policies for bilingual and trilingual cantons (*Rash, 2002*), in conjunction with the relatively low number of speakers, about 7.6% (*worldatlas.com, 2002*). Romansch is spoken by 0.6% of Swiss, but is not recognised by AltaVista, along with many other minority languages in the EU.

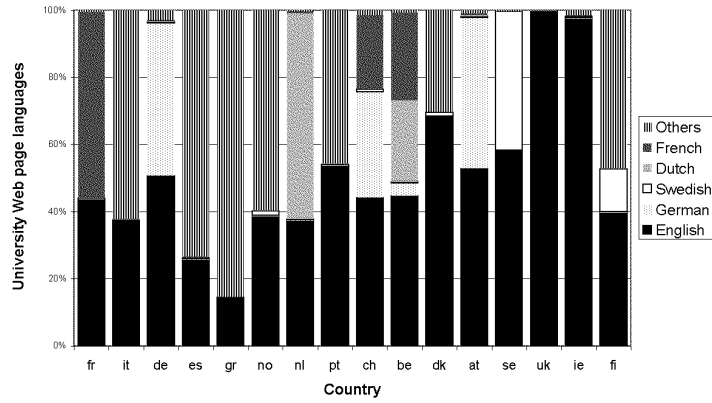


Figure 2. Language breakdown of the Web pages hosted by the universities in the large Western European countries. English is at the bottom of the graph. Only significantly multinational languages are shown. For all countries with “other” language above 20% of the total, it represents almost exclusively the dominant language in the country

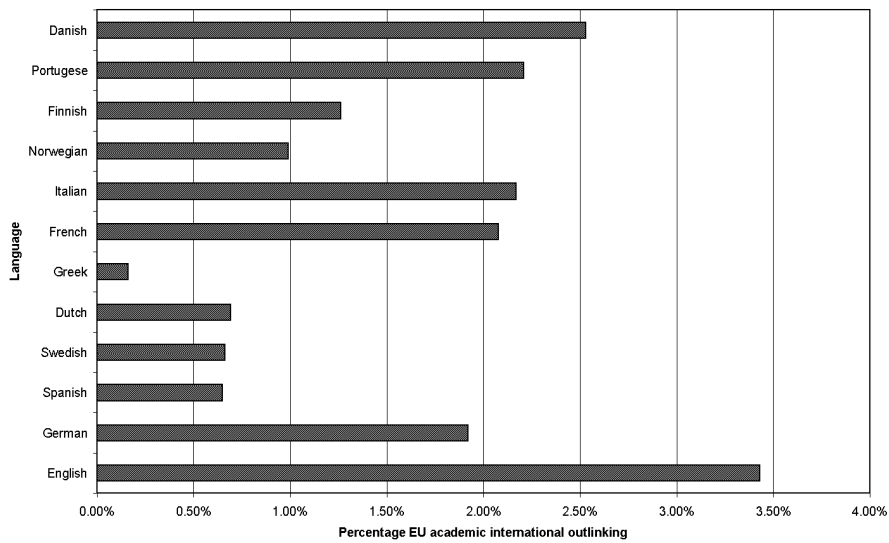
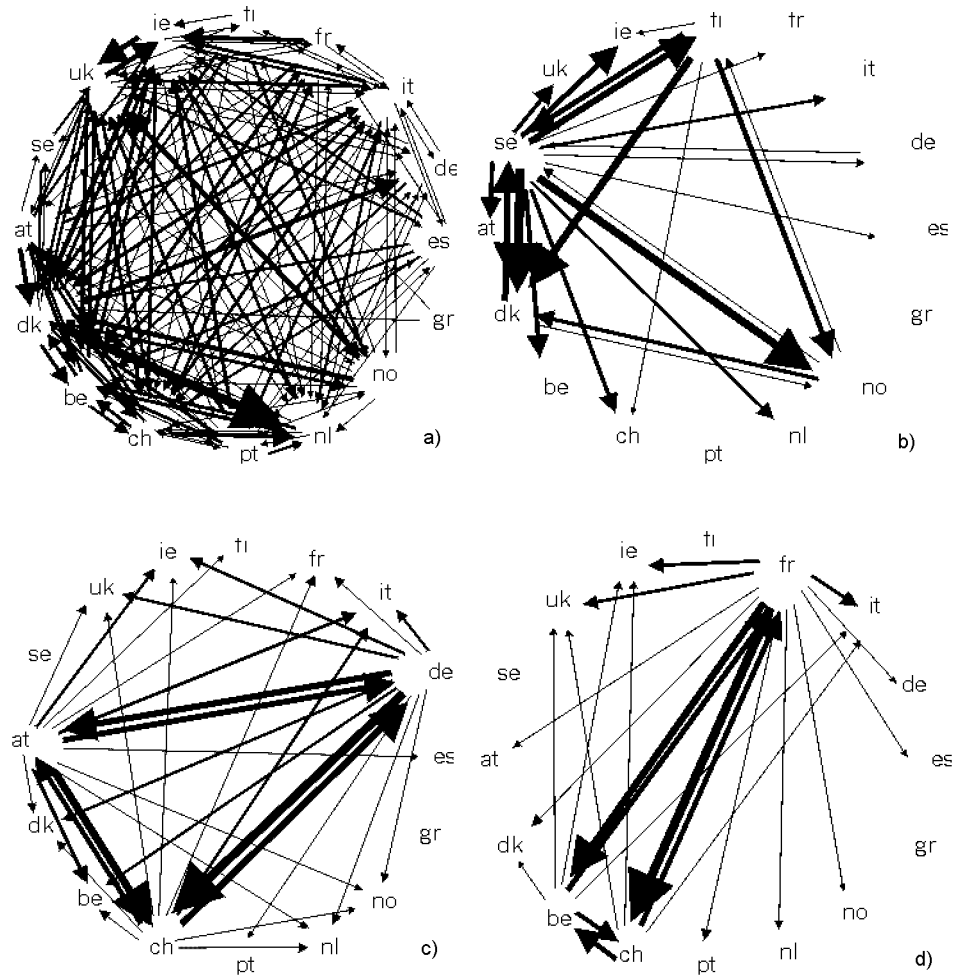
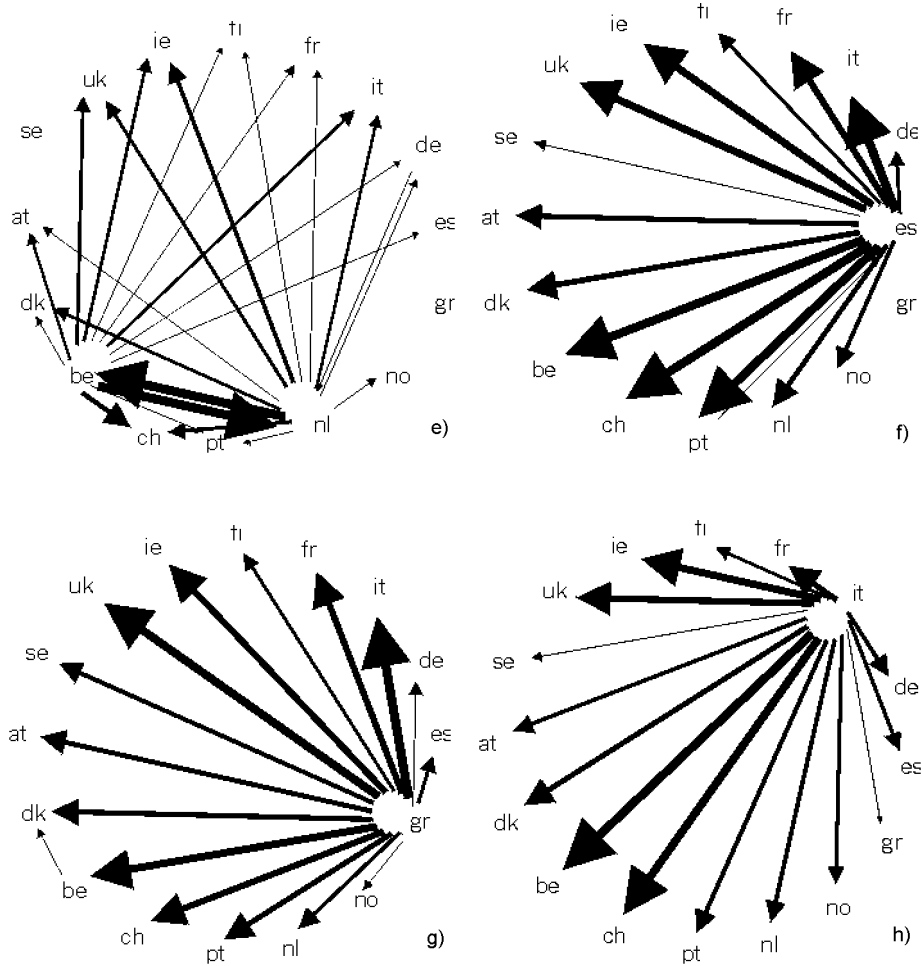


Figure 3. The percentage of the total number of pages in each language (from all countries combined) that host international outlinks to other universities in the study. Languages are displayed in order of total Web page counts





Figures 4. Normalised linking between EU countries by source page language. a) English, b) Swedish, c) German, d) French, e) Dutch, f) Spanish, g) Greek, h) Italian. Arrow widths are normalised for the total number of Web pages (in all languages) in both source and target university system. They are also normalised to make the thickest arrow in each diagram of the same size and so should be viewed in conjunction with Figure 1 for the relative size of each language

From Figure 3 it can be seen that there are considerable differences in the extent to which different language pages in all countries host international outlinks. It can be seen that English is the most international language in this respect and Greek the least. Since some of the languages have a significant presence in more than one country, this probably represents overlapping linguistic and national trends. We have not attempted to separate these two factors in this diagram, but the issue is addressed in later figures.

Linguistic patterns of web interlinking

Figures 4a-h give patterns of interlinking along linguistic lines. All the European languages not shown have graphs similar to that for Italian: i.e. with arrows all originating at their linguistic home country.

Several points stand out as a result of inspection of the diagrams. Firstly, it is clear that some languages are significantly multinational.

- *English* is widely used for link source pages throughout Europe, without a significant tendency to link to or from the UK in this language. It may be concluded that English is a standard Web language for linking throughout the EU. It can be seen, however, that Greece is less extensively interlinked in English.
- *German, French and Dutch* are all official languages in more than one EU country, and this is clearly reflected in the diagrams. Note that variations of German are official languages also in Austria and Switzerland, French is an official language of Belgium and Switzerland and Dutch is an official language of Belgium.
- *Italian* is also an official language of Switzerland, but is only spoken by about 7.6% of the population (*worldatlas.com*, 2002), which may account for the insignificant linking from Switzerland to Italy.
- *Swedish* is an official language of Finland as well as Sweden but its graph shows that it is extensively used for international link pages between all pairs of countries in Scandinavia (except from Denmark to Finland). Note that the patterns of interlinking for the other Scandinavian languages of Finnish, Norwegian and Danish are similar to that of the Italian graph, in that all significant linking takes place only from the main country. International interlinking in Scandinavia appears to take place predominantly in the native languages of the source country or Swedish. Swedish is also used to link between other pairs of countries, neither of which have it as an official language. This is the only language covered apart from English that has this

extensive multinational character outside of its official countries. As with all the languages covered, it is unclear the extent to which the international linking represents multilingualism or whether the links are sourced by, or targeted at, native speakers in countries where their language is not recognised as official.

There are also a number of anomalies on the graphs: links that violate the general patterns that are evident or have been discussed above.

- *Spanish*: Portugal hosts a significant number of pages in Spanish that link to its neighbour Spain.
- *Greek*: Belgium hosts two links to Denmark, but this number was large enough to warrant an arrow due to the low overall international Greek interlinking and the small size of both countries. The links were from two pages of a multilingual site for an EU funded project Eurotroph, investigating the functioning of coastal ecosystems, which did not include a Greek partner <http://www.ulg.ac.be/oceanbio/eurotroph/public/index.htm>.
- *English*: The biggest normalised link direction is from Denmark to Holland. This was pursued as a potential anomaly and the major cause was identified as a collection of Web pages that were created from email discussions on topics in computing (Svavarsson, 2002). The discussions had international contributors and the links were actually hyperlinked email addresses. This must be classified as an anomaly because of the relatively artificial nature of the links and the fact that the content of the material hosted on the site was not created by its owner.

Interpretation of results

The results presented so far have assumed that Web links are a form of scholarly communication and that it is meaningful to count them up and form conclusions based upon these totals. As discussed earlier, the former assumption is based upon a study showing that this is the case for links between UK universities whereas the latter is based upon correlations found between counts of links and research indicators in the UK, Australia and Taiwan. Nevertheless, it is still worth discussing the factors that must be taken into account when interpreting the results. Firstly, there are a number of logical reasons that may account for a large amount of interconnectivity between university Web sites in different countries and a specific language.

1. Scholars in country A publish material in language X in the belief that this is the language that their target audience will be able to read. This would

particularly apply to indigenous languages, English and Swedish (in Scandinavia).

2. A Web site in country A hosts a mirror site of pages in language X from country B, which contains many links back to the creating university.
3. Many scholars from country A move to country B to study or teach and publish in their own language and/or link back to their home country.
4. Individual sites contain huge numbers of repeated international links. The Alternative Document Models (Thelwall, 2002) were designed to combat this but cannot be applied to data derived from search engines.
5. The scholars in country A are able to read the pages in country B due to sharing a common language and so link to them more frequently.
6. Geo-political factors make scholars in country A particularly aware of activities in country B.
7. Individual large collections of international links. For example a list of all German language media Web sites hosted by a UK German studies department.
8. International regional research collaboration initiatives, perhaps particularly in evidence for Scandinavia and Benelux. Examples are the Nordic Federation of Research Libraries Associations, and the Benelux Yeast Research Groups.

All of these hypotheses are difficult to test individually. Probably they all apply somewhere in the data set, which complicates significantly the task of interpreting results. Bibliometrics has a similar problem with the interpretation of citation counts but Web links are much more varied in scope. The three potentially most damaging causes are (2), (4) and (7). These three are particularly threatening because they are capable of producing high counts without the existence of a high underlying degree of international awareness. As a result, the main conclusions were tested by sampling the Web pages containing the links in order to check that (2), (4) and (7) are not dominant causes.

- There is a high degree of international interlinking in English.
- There is a high degree of international interlinking in Swedish.
- There is a high degree of international interlinking between countries sharing a common language.

These were tested by viewing a random sample of AltaVista results pages, ten in each case. One anomaly had already been found, between Denmark and the Netherlands, as mentioned above. Some examples of types of multiple pages from a common source in English are given below.

- *Mirror sites* The Linux Gazette mirror copy in France, linking to Norway in English, http://ftp.iut-bm.univ-fcomte.fr/LINUX/lg/www_root/issue20/lyx.html
- *International sites* An international trilingual resource for mathematicians interested in Navier-Stokes equations: <http://wwwlma.univ-bpclermont.fr/NSenet/>
- *Repeated international links* Franz J. Hauck maintains identical links to an associated research group in the Netherlands from several of his pages, e.g. <http://www4.informatik.uni-erlangen.de/~fzhauck/Admin/>

In no case investigated did the repeated links dominate the standard links, but the investigation of English raised the likelihood that the figures for English in particular are inflated by what seems likely to be an above average incidence of type (2) and (7) anomalies. It is accepted and this is a rather cursory investigation of this important point, but its complexity speaks for a series of future more detailed investigations.

For Swedish, the high linking from Germany was traced to a Nordic literature online initiative mirror site of Swedish biological information in Germany <http://www.biologie.uni-hamburg.de/b-online/lindman/>. Linking between Scandinavian countries in Swedish did include anomalies. For example, several links from Norway to Finland in Swedish came from the NordLingNet online mailing list <http://www.hd.uib.no/nordlingnet/>. Anomalies were also sought in linking between Belgium and the Netherlands in Dutch, but none were found.

Conclusion

It is clear from the data that English is the dominant language of the Western Europe academic Web and an important language for the aspect of international scholarly communication that Web links represent. The majority of pages were in English, a higher proportion of English language pages contained international European links than any other major language and all countries except Greece were extensively linked to and from in English. This answers the first research question.

The second question is actually also answered by the case of English. Apart from Greece there is extensive interlinking throughout Western Europe, both in host country native languages and English. It was also clear that linking between countries with a shared language was higher in academia, but this was far from being the only type of linking present. Swedish was also confirmed to be an international language in Scandinavia.

Returning to the case of Greece, there is strong evidence of its isolation from Western Europe. No country links significantly to it in English and neither do most in any other language, with the exception of its neighbour Italy. The graph for linking in Greek shows that Italy is its preferred choice, but all linking is relatively weak as a consequence of Greek pages being the least likely to link externally. Of course Greece may well link more extensively to Eastern European countries, some of which are nearer to it and are linguistically closer. The two non-EU states did not stand out as unusual in the diagrams and so the formal EU grouping does not appear to have significantly impacted on the patterns of online informal scholarly communication yet. This may be because Norway and Switzerland have been permitted to participate in EU funding devoted to promoting collaborative scholarly research, including ESPRIT (ESPRIT, 2002) and the EU's Fifth Framework Programme (CORDIS, 2002).

In summary, the majority of the Western European academic Web is international and multilingual in character, with English and national languages operating in tandem throughout, with Swedish also in Scandinavia.

*

Many thanks to the referees for helpful comments and corrections.

References

- ALMIND, T. C., INGWERSEN, P. (1997), Informetric analyses on the world wide web: methodological approaches to 'webometrics', *Journal of Documentation*, 53 (4):404–426.
- BAR-ILAN, J. (1999). Search engine results over time – A case study on search engine stability, *Cybermetrics*, 2/3. Available: <http://www.cindoc.csic.es/cybermetrics/articles/v2i1p1.html>
- BJÖRNEBORN, L., INGWERSEN, P. (2001), Perspectives of webometrics, *Scientometrics*, 50 (1): 65-82.
- BORGMAN, C., FURNER, J. (2002). Scholarly communication and bibliometrics. In: CRONIN, B. (Ed.), *Annual Review of Information Science and Technology*, 36, Medford, NJ: Information Today Inc. pp. 3-72.
- BRIN, S., PAGE, L. (1998), The anatomy of a large scale hypertextual web search engine, *Computer Networks and ISDN Systems*, 30 (1-7): 107–117.
- CORDIS (2002), *Mini Guide To The Fifth Framework Programme (FP5)*, Available: http://www.cordis.lu/united_kingdom/fp5-mini.htm
- CRONIN, B. (2001), Bibliometrics and beyond: Some thoughts on web-based citation analysis, *Journal of Information Science*, 27 (1): 1–7.
- CRYSTAL, D. (1997), *English as a Global Language*, Cambridge: Cambridge University Press.
- CRYSTAL, D. (2001), *Language and the Internet*, Cambridge: Cambridge University Press.
- DAVENPORT, E., CRONIN, B. (2000), The citation network as a prototype for representing trust in virtual environments, In: CRONIN, B., ATKINS, H. B. (Eds), *The Web of Knowledge: A Festschrift in Honor of Eugene Garfield*, Metford, NJ: Information Today Inc. ASIS Monograph Series, pp. 517–534.
- ESPRIT (2002), *Non-EU Participation in ESPRIT*, Available: <http://www.cordis.lu/esprit/src/non-eu.htm>

- EUROPA (2001), *European Commission – Research: Fifth Framework Programme*, Available: <http://europa.eu.int/comm/research/fp5.html>. Accessed 11 July 2002.
- EUROPA (2002a), *The Sixth Framework Programme (2002-2006)*, Available http://europa.eu.int/comm/research/fp6/index_en.html. Accessed 11 July 2002.
- EUROPA (2002b), *The Member States of the European Union*, Available http://europa.eu.int/abc/eu_members/index_en.htm. Accessed 17 July 2002.
- FLOWERDEW, J., PEACOCK, M. (2001), *Research Perspectives on English for Academic Purposes*, Cambridge: Cambridge University Press.
- GARFIELD, E. (1967), English – An international language for science? *Current Contents*, Dec 26, 19–20.
- GARRIDO, M., HALAVAIS, A. (2002), Mapping networks of support for the Zapatista movement: Applying social network analysis to study contemporary social movements, In: M. MCCAUGHEY, M. AYERS (Eds), *Cyberactivism: Critical Practices and Theories of Online Activism*, London: Routledge.
- GLÄNZEL, W., SCHUBERT, A. (2001), Double effort = Double impact? A critical view at international co-authorship in chemistry, *Scientometrics*, 50 (2): 199–214.
- HOFFMANN, C. (2000), The spread of English and the growth of bilingualism with English in Europe, In: J. CENOZ, U. JESSNER (Eds), *English in Europe: the Acquisition of a Third Language*, Clevedon: Multilingual Matters, pp. 1–29.
- JENKINS, J. (2002), A sociolinguistically based, empirically researched pronunciation syllabus for English as an international language, *Applied Linguistics*, 23 (1): 83–103.
- JIMÉNEZ-CONTRERAS, E., LÓPEZ-CÓZAR, E. D., RUIZ-PÉREZ, R., FERNÁNDEZ, V. M. (2002), Impact-factor rewards affect Spanish research, *Nature*, 417: 898.
- KLEINBERG, J. (1999), Authoritative sources in a hyperlinked environment, *Journal of the ACM*, 46 (5): 604–632.
- LAWRENCE, S., GILES, C. L. (1999), Accessibility of information on the web, *Nature*, 400: 107–109.
- MOED, H. F. (2002), Measuring China's research performance using the Science Citation Index, *Scientometrics*, 53 (3): 281–296.
- PARK, H. W., BARNETT, G. A., NAM, I. (2002), Hyperlink-affiliation network structure of top web sites: Examining affiliates with hyperlink in Korea, *Journal of the American Society for Information Science*, 53 (7): 592–601.
- PENNYCOOK, A. (1996), *The Cultural Politics of English as an International Language*, London: Addison-Wesley Pub. Co.
- POLANCO, X., BOUDOURIDES, M. A., BESAGNI D., ROCHE, I. (2001), *Clustering and Mapping Web Sites for Displaying Implicit Associations and Visualising Networks*, University of Patras, Greece.
- RASH F. (2002), The German-Romance language borders in Switzerland, *Journal of Multilingual and Multicultural Development*, 23 (1): 112–136.
- ROUSSEAU, R. (1999), Daily time series of common single word searches in AltaVista and NorthernLight, *Cybermetrics*, 2/3. Available: <http://www.cindoc.csic.es/cybermetrics/articles/v2i1p2.html>
- SKUTNAB-KANGAS, T. (2000), *Linguistic Genocide in Education or Worldwide Diversity of Human Rights?* New York: Lawrence Erlbaum Associates.
- SMITH, A., THELWALL, M. (2002), Web impact factors for Australasian universities, *Scientometrics*, 54 (3): 363–380.
- SVAVARSSON, M. (2002), Index of /~magnus/MHonArc. Available: <http://www.sprog.auc.dk/~magnus/MHonArc/>
- THELWALL, M. (2001a), The responsiveness of search engine indexes, *Cybermetrics*, 5 (1) <http://www.cindoc.csic.es/cybermetrics/articles/v5i1p1.html>
- THELWALL, M. (2001b), Extracting macroscopic information from web links, *Journal of the American Society for Information Science and Technology*, 52 (13): 1157–1168.

- THELWALL, M. (2001c), Exploring the link structure of the Web with network diagrams, *Journal of Information Science*, 27 (6):393–402.
- THELWALL, M. (2002a), Conceptualizing documentation on the Web: an evaluation of different heuristic-based models for counting links between university web sites, *Journal of the American Society for Information Science and Technology*, 53 (12):995–1005.
- THELWALL, M. (2002b), Methodologies for crawler based web surveys, *Internet Research: Electronic Networking and Applications*, 12 (2):124–138.
- THELWALL, M., BINNS, R., HARRIES, G., PAGE-KENNEDY, T., PRICE E., WILKINSON, D. (2002), European Union associated university websites, *Scientometrics*, 53 (1):95–111.
- THELWALL, M., SMITH, A. (2002), A study of the interlinking between Asia-Pacific University Web sites, *Scientometrics*, 55 (3):335–348.
- THELWALL, M., TANG, R. (2002), *Disciplinary and Linguistic Considerations for Academic Web Linking: An Exploratory Hyperlink Mediated Study with Mainland China and Taiwan*, University of Wolverhampton.
- TREFFERS-DALLER, J. (2002), Language use and language contact in Brussels, *Journal of Multilingual and Multicultural Development*, 23 (1):50–64.
- VAN RAAN, A. F. J. (2001), Bibliometrics and Internet: Some observations and expectations, *Scientometrics*, 50 (1):59–63.
- VAUGHAN, L., HYSEN, K. (2002, to appear), Do Web link counts resemble citation counts: An empirical examination, *ASLIB Proceedings*.
- VAUGHAN, L., THELWALL, M. (2003, to appear), Scholarly use of the Web: What are the key inducers of links to journal Web sites? *Journal of the American Society for Information Science and Technology*.
- WILKINSON, D., HARRIES, G., THELWALL, M., PRICE, E. (2003, to appear), Causes of academic Web site interlinking: Evidence for the Web as a novel source of information on informal scholarly communication, *Journal of Information Science*, 29 (1) 59–66.
- WILLEMYNS R. (2002), The Dutch-French language border in Belgium, *Journal of Multilingual and Multicultural Development*, 23 (1):36–49.
- WORLDATLAS.COM (2002), *World Atlas Switzerland, Europe, Information Page*. Available: <http://www.worldatlas.com/webimage/countrys/europe/chcia.htm>
- ZITT, M., BASSECOULARD, E., OKUBO, Y. (2000), Shadows of the past in international cooperation: Collaboration profiles of the top five producers of science, *Scientometrics*, 47 (3):627–657.