

Chapter IV

Foreign Language Analysis

The most difficult area to analyze in academic library collections is that of foreign language holdings. While it is a reasonable assumption that the majority of English language titles are cataloged relatively quickly, that is, within a few years after publication and acquisition, it has been suggested through a number of studies that foreign language materials are acquired and cataloged over a much longer span of time.

Groverⁱ conducted studies on the speed and availability of copy cataloging for Latin American books. Grover searched a sample of 298 humanities and social sciences Latin American Spanish language approval monographs in both OCLC and RLIN databases from June to August 1983 to determine the speed of copy cataloging. Upon receipt 37% had copy on OCLC. A year after receipt, 50% had not been cataloged anywhere in the U.S. After 18 months 88% had cataloging records. One-third was cataloged in the first six months and one-twelfth was cataloged in the next six months. Both utilities had the same number of books, but not the same titles.

Sercanⁱⁱ carried out a similar study at Cornell for 783 Latin American Spanish language books received on blanket order. Upon receipt full cataloging was available for only 4% of the books. RLIN had catalog records for 19%. In process records were found for 60% of the titles. The titles were searched every four weeks. Sercan found that allowing more time slightly improved the number of titles found, but there was a marked decline in copy cataloging from the findings of previous studies.

Gurevichⁱⁱⁱ searched 507 Russian-language monographs received at the University of Texas-Austin from December 5, 1989-May 5, 1990. Monographs older than 3 years were

excluded so that the study concentrated on new or current publications. Bibliographic records were found for 63% on the titles in OCLC; member copy made up a large portion, as 65% of the records lacked LC call numbers.

Tsao^{iv} searched 518 Chinese and 250 Japanese books received at the University of Chicago Library from November 1992 to March 1993. Fifty percent of the Chinese and 53% of the Japanese titles were in RLIN. A later search found few additional records for the Chinese titles, but the hit rate for the Japanese books improved to 76%. As with other studies, LC copy was far exceeded by member copy, six times for the Chinese and twice as many member records for the Japanese titles.

The majority of these studies were conducted with blanket order or approval plan receipts. Thus the hit rates in these studies are probably higher than would obtain for a broader range of materials. Seymour searched for OCLC copy on 431 East European monographs received from August to December 1993 at the Hoover Institution Library and found copy for only 20%. Of the 85 cataloging records found, 39% were from the Library of Congress.^{p.275} Because of the depth of collecting at the Hoover Institution Library these publications are probably more representative of special collecting. These findings could be indicative of the lack of cataloging copy for any but the more mainstream materials.^v

Cataloging lag for foreign imprints has also been indicated through data on foreign imprints first analyzed in the ARL/AAU study, *Scholarship, Research Libraries, and Global Publishing* (1996), and updated in 1999. The ARL data are for foreign *imprints* not foreign *language* imprints and are not comparable to data extracted from the OCLC/AMIOGOS CACD. There can be considerable numbers of English language titles published in Europe, Africa, India, and Australia included in the foreign imprint figures. The ARL data include foreign libraries which are OCLC members and not just North American academic libraries as in the CACD, but

the cataloging trends are for bibliographic records on foreign titles available through WorldCat. Another aspect of the ARL WorldCAT data is that the number of libraries is not held constant as in the CACD, so that the data for different imprint years can represent varying numbers of libraries' holdings. The ARL data from WORLDCAT for 1994 foreign imprints indicate that over 50% of the total number of 1994 imprints were cataloged between 1994 and 1999. In 1999, the number of 1988 imprints cataloged was near 15% of the total number of 1988 foreign imprints in World Cat^{vi}. From these data we can see that cataloging data on numbers of foreign imprints only become reliable five years or more after the publication date.

With this caveat in mind, the data from the 1997 OCLC/AMIGOS CACD are utilized in this chapter to examine the profile of foreign language holdings in the three peer groups of academic libraries and the CACD database for the 1987-1995 time frame. The analysis does not use the data for 1994 and 1995; 1993 data are four years from the production of the 1997 CACD. The average number of years between publication and entry of records and holdings in the bibliographic utilities for foreign language monographs has not been established. The data in this study indicate that the number of foreign language imprints rose each year from 1987-1993, after which the total number of titles begins to decrease. The analyses in this chapter concentrate on this shorter time span in which data appear to be more reliable for foreign language imprints.

In chapter III, nine years of data for the three peer groups of academic libraries were analyzed by total numbers of imprints according to percentage of total by subject, unique titles, mean number of holding libraries, and percentage of 1-5 holding libraries. In this chapter, the ratio of English language to non-English language titles and seven foreign language groupings are analyzed. Complete data tables for all seven language groupings, plus English and non-English language groupings for the four peer groups are contained in Master Table Series One-Three on the ACRL website. Summary tables drawn from those data are contained in this

chapter. The first analysis looks at the total number of titles divided into two major divisions of English and Non-English language imprints.

Ratios of English/Non-English Language Titles

The ratios of English language monographs to non-English language monographs for the nine-year period are displayed in Table IV-1. For PG1, the ARL libraries, it can be seen that the ratio shifts slightly each year in favor of English language materials. From a very close 50.37% English to 49.63% non-English in 1987, the ratio shifts to 51.53% English and 48.47% non-English in 1993. Although not directly comparable because of a different number of libraries in the datasets, previous studies of the aggregated collections of ARL libraries found a nearly perfect 50/50 ratio in 1985.^{vii} This is the only baseline available for comparison prior to the time period of the current study. The nine-year dataset here shows that the shift toward English language materials continued throughout the late 1980s and the first half of the 1990s.

In both PG4 and PG7 the ratio of English to non-English is predominantly in favor of English throughout the time span. In PG4 it begins with a 76.5% English to 23.5% non-English in 1987. The percentage of English language monographs does increase to 79.8% by 1993. This is the largest difference, 3.3 percentage points, between the 1987-1993 ratios of any of the peer groups. The decrease in the number of non-English language titles in PG4 seems to be a pattern of decline even though the percentage decline is slight. The difference in the ratio occurs because the total number of titles fluctuates and the number English language titles follows the pattern of the total.

In PG7, as in PG4, there is little fluctuation in the number of non-English language titles. The biggest change is in the decrease in the number of English language titles in 1988 and 1989 from 1987. In 1990 the percentage of English language titles returns to the level of 1987 and increases in 1992 and 1993. In 1987 the ratio is 87.3% English to 12.7 % non-English and in 1993 the ratio is 88% English to 12% non-English. Although there is some fluctuation in the years in between, the ratio is the same in 1993 as in 1987 for PG7.

Table IV-1

Ratio of English to Non-English Titles

The ratios for both PG4 and PG7 show that English is the predominant collecting language in those peer groups. It is quite likely that these proportions have not been influenced by the economic factors which affected the ARL collections, but that these ratios for non-ARL academic libraries reflect collecting patterns which have obtained all along.

The most interesting observation for peer groups 4 and 7 is not in the ratios which remain relatively consistent within each group, but that the absolute numbers reveal an interesting pattern. The absolute numbers for English language imprints in both peer groups are not far apart, staying within less than 5,000 titles difference in most years. The difference in the number of non-English imprints, however, is much larger. PG4 has roughly twice the number of non-English imprints as PG7, with this ratio remaining relatively constant across the nine-year period. Thus, it is the number of foreign language imprints that are determining the difference in the ratio of English to non-English in both peer groups with the number of English language imprints in both groups being very similar. PG4 does contain libraries in more research oriented institutions than those in PG7 and this difference shows up sharply in the much larger number of foreign

language imprints in PG4.

In PG14, we would naturally expect that English language materials would predominate given that public libraries are also included as well as special and smaller academic libraries. From a nearly 60/40 ratio in 1987, the percentage of English language materials actually declines to 57.3% English and 42.6% non-English in 1993. But the ratio is clearly weighted toward English language materials. The “weight” of influence of the ARL collections can be seen in the aggregated database of all libraries in the 1997 CACD in PG14. The ratios of English to non-English are different in the non-ARL academic libraries, but those collections do not contribute as many titles to the database as do the ARL libraries. As in many of the variables analyzed in this study, PG1 and PG14 show nearly identical collecting patterns and trends with respect to the ratio of English to non-English language titles. Although the ratios for English/non-English are not exactly the same, the percentage ranges are similar in PG1 and PG14 showing the influence of the ARL library group in the aggregated database.

It is clear from Table IV-1 that the two non-ARL academic library peer groups have a much lower level of collecting of foreign language materials in comparison with the ARL group. While the collecting patterns for PG4 and PG7 differ from each other, the ratios are much higher for English language titles in those peer groups than in PG1 and PG14.

Again, it is not known what effect cataloging lag has on the number of titles in the dataset in 1994 and 1995 but, in 1994, almost 3 years prior to the construction of the 1997 CACD, the numbers begin to decline for English language materials as well as foreign imprints. As the numbers reported in both the ARL and ACRL statistical series for monograph acquisitions also decline in 1993, and only show slight improvement in 1995, the numbers in the CACD dataset appear to be a real decline in purchasing in 1993-1995 for English language materials. Foreign language materials are still subject to both later acquisitions and cataloging lag so that the numbers are still less reliable for these monographs.

As was found in the analyses in chapter III for total number of titles by peer group and subject groupings, the data exhibit remarkable consistency in the ratio of English language to non-English language titles across the nine years, even though the absolute numbers may vary.

Thus far, the analysis has looked at the total number of titles by peer group and by ratio of English to non-English language titles. The next analysis using all four peer groups is for the number of titles and proportions of titles by seven foreign language groupings.

Foreign Language Groupings as a Percentage of Total

Seven major foreign language groupings are analyzed in this study. The data for the language groupings are extracted for the CACD by utilizing the fixed field tag in the MARC record for language. The seven language groupings analyzed in the study are Arabic, Chinese, French, German, Japanese, Russian, and Spanish.

Table IV-2 shows the seven foreign language groupings along with totals for English and non-English language groupings as percentages of all imprints for each year. It is obvious from the absolute numbers in all of the peer groups that the rankings from most to least collected, with few exceptions, are German, Spanish, French, Russian, Chinese, Japanese, and Arabic. The absolute numbers and percentages of total for the seven languages do not add up to the percentage for non-English because other languages not analyzed in the study would be included in the non-English category. But it can be seen that the seven languages analyzed do add up to a majority of the non-English titles. In PG14, the seven languages make up 64% of the non-English category. Likewise, the percentage of total is 67% in PG1, with 82% in PG4 and 84% in PG7.

The patterns of collecting are similar to those which have been established in prior analyses in this study. Again, PG14, the database, and PG1, the ARL libraries are very close in absolute numbers and percentage of total for the seven languages. PG4 has far fewer absolute

numbers in the seven languages with PG7 having an even lower level of collecting activity.

From these data it would seem that if cataloging lag is not a major factor for the years 1987-1993 the number of titles would decrease moving forward but that is not the pattern here. From 1987-1994 the numbers of non-English titles increase in PG4 and PG7. Beginning in 1990, the numbers increase in PG14 and PG1. This would seem to corroborate earlier findings that the acquisition of foreign language titles did decrease in the ARL libraries in the latter 1980s.^{viii}

Table IV-2

Foreign Language Groupings as a Percentage of Total

The total number of imprints, both English and non-English, decrease in 1995 in all peer groups. For both PG14 and PG1, the total number of titles increases each year through 1992, remains relatively steady in 1993 and then begins to decrease in 1994. In the non-ARL academic library peer groups, the same pattern obtains, only with a flatter trend line that is, the increases are not as marked as in the larger peer group

Within the overall trends, the seven languages do have slightly differing patterns. In PG14, the peak years in total are 1992 and 1993. The seven languages also have their highest number of titles in those years. If cataloging lag can be assumed to be in the neighborhood of five years with the CACD constructed in 1997, then these data show that collecting activity for the years 1987-1991 was lower than in the later years of 1992-1993. From Table IV-2 we can see that in both of the large peer groups, the number of titles for German and Spanish are lower in 1987 than in 1993. Whereas French has a slightly higher total in the ARL in 1993 than in 1987, in PG14, French has virtually the same number from 1987 through 1994. Russian begins to decline in numbers after 1991 in both peer groups. The numbers for the lesser collected

languages are very close in both peer groups. Arabic does show an increase in PG14 moving forward, while it decreases in the ARL libraries. Japanese displays a similar pattern.

In PG4, French and German both decline going toward the present, while Spanish and Russian hold steady. The absolute numbers for all languages are much lower than in the larger peer groups. In PG7 the numbers for French and German are steady, while Spanish increases slightly. In both PG4 and PG7, the other languages show negligible collecting.

The next analysis shows the concentration of foreign language collecting by broad knowledge groupings.

Percent of Broad Subject Groupings: English and Non-English

The percentage of total for the broad subject groupings of humanities, social sciences, and sciences are shown in Table IV-3 along with the ratio of English to non-English titles within each knowledge grouping. The ratios are quite different for the three knowledge groupings. In 1987, in PG1 the humanities actually have a larger proportion in foreign language titles than in English language titles. The ratio of English to non-English is 35/65. The social sciences have a 51/49 English to non-English ratio, and the sciences have an 80/20 ratio. In the humanities, English increases nearly two percentage points by 1993, but remains nearly the same in the social sciences. The largest difference occurs in the sciences in which English language titles increase five percentage points between 1987 and 1993. Or phrased in the reverse, non-English language titles in the sciences decrease five percentage points. The stability of the ratios in the social sciences would seem to indicate that collecting patterns remained stable and that cataloging lag was not much of a factor until 1994. The changes in the ratio over time in the sciences may indicate a real decline in the collecting of foreign language monographs in the sciences from 1987-1993. Given the pressures on the monograph budgets occasioned by the escalation in serials prices which were most severe in the sciences, it would seem logical that foreign language collecting would be retrenched in the sciences to allow continued purchasing in the humanities and social sciences. Cataloging lag could be a factor in the appearance of a decline in foreign language monographs in the sciences, in that those materials might be seen to be in less demand than foreign language materials in the humanities and social sciences and hence not be a cataloging priority. But given the economic conditions, the decline would seem to be real.

Table IV-3

Percent of Broad Subject Groupings: English and Non-English

In PG 4 the humanities ratio is reversed from that of PG1. English language titles make up 60% of Humanities titles with non-English at 40 percent. The percentages are almost exactly 10 percentage points different for PG 7; English makes up 71% and non-English 29%. As in PG1, the percentage of English language monographs increases: 5 percentage points by 1993 in both PG4 and PG7. The pattern is similar, only the percentages vary between peer groups.

In the social sciences in PG4 the ratio is 80/20 in favor of English language titles. The ratio alters slightly to 83/17 by 1993. The percentage of English in the social sciences in PG7 is much higher at 90/10, varying as in PG4 to 2 percentage points higher for English in 1993. Again, although the ratios are different in PG4 and PG7, they follow a similar pattern.

The ratios for the sciences are very close to the same in both PG4 and PG7. Neither peer group is collecting foreign language monographs in the sciences. The percentage English language does not drop below 95% in either peer group. In 1993 the percentage of foreign language titles is 3% in PG4 and less than 2% in PG7 in the sciences.

As can be seen in Table IV-3, PG14 and PG1 have almost identical percentages of total for the English/non-English ratios in the humanities and sciences. But the ratios differ between the two peer groups in the social sciences. There is a larger percentage of English language titles in PG14 in the social sciences than in PG1. The preponderance of English language titles in the smaller academic library peer groups is an influence in the social sciences. The extent to which the foreign language holdings of the ARL are also the foreign language holdings of the CACD database is made very apparent in Table IV-3.

Differences in collecting intensity among the language groupings are shown by the variable of mean number of holding libraries in the next section.

Mean Number of Holding Libraries by Language

It was observed in chapter III that the mean number of holding libraries appears to be less affected by cataloging lag than the number of unique titles in that once a bibliographic record is available for a title, library holdings symbols can be quickly added by other libraries. The intensity of collecting shows more dramatically in the mean number of holding libraries by language grouping. Table IV-4 displays summary data from the four library peer groups for the mean number of holding libraries by all titles, the divisions of English/non-English, and the seven language groupings in the study. Predictably, the mean number of holding libraries is much higher for English language titles than for non-English titles. In the previous section, it was established that in PG1, the ratio of English to non-English titles is approximately 52% to 48%. For English language monographs, the mean number of holding libraries is 17.71, while non-English titles have 5.80 holding libraries in PG1. English language titles have three times as many holding libraries as non-English titles on the average.

Bearing in mind that the number of libraries in the peer group are reflected in the mean number of holding libraries, the data in table IV-4 are even more remarkable in that the mean number of holding libraries for the paired peer groups are close in absolute numbers. In peer groups 14 and 1, the mean number of holding libraries for all non-English titles are less than one percentage point different. The collecting of French, German, and Spanish in peer groups 4 and 7 would seem to be raising the average for those languages in PG14. In the languages not being collected in peer groups 4 and 7, the mean numbers for peer groups 14 and 1 are very close. Thus, from table IV-4, it can easily be seen that the ARL libraries are the chief contributors to the mean numbers for foreign language titles in peer group 14. It can also be seen that, although collecting is at a much lower level of intensity in peer groups 4 and 7, even those low mean numbers seem to have an influence on the averages for the database, PG14. While mean numbers

in the 6-8 range for number of holding libraries per title might not seem low for the 95 libraries in the ARL group, those same mean numbers for the 2,646 libraries in PG14 do seem shockingly low. It is easily understood that the means for PG4 and 7 are very similar in that the libraries in PG7 are in institutions granting fewer graduate degrees than those in PG4. Thus, although PG7 has more libraries, the libraries in PG4 are larger in size and are components of larger universities.

Another aspect that is shown in Table IV-4 is that the mean number of holding libraries for English language monographs is very near the same percentage of the number of libraries in the peer group for PG1 and PG4. In PG1, the mean number of 17.71 translates into 18.7% of the libraries on the average owning a title. In PG4, the 20.65 mean number of holding libraries is 16.8% of the total number of libraries in the peer group. But in PG7, 15.87 for the mean number of holding libraries is only 7% of the total number of libraries. The indication would seem to be that in PG7 there is far less agreement on the acquisition of the same titles than in the larger academic libraries.

It is not possible to determine from the mean number of holding libraries per title by peer group data in Table IV-4 if the collecting of foreign language materials in the academic library peer groups is systematic collecting by a small number of libraries, or much more dispersed collecting by a larger number of libraries within the groups. It would appear that for PG4 and PG7, the collecting is by a small number of libraries. Within the ARL the same is probably true with wider collecting in only major fields or disciplines.

Table IV-4

Mean Number of Holding Libraries by Language Grouping

	Peer Group 14	Peer Group 1	Peer Group 4	Peer Group 7
No. of Libraries	2,646	95	123	227
All Languages	40.11	11.99	12.27	14.12
English	64.91	17.71	20.65	15.87
Non-English	6.68	5.80	2.42	2.13
Arabic	5.45	5.23	1.13	1.38
Chinese	4.27	3.92	1.28	1.23
French	8.81	7.23	2.70	1.99
German	9.23	7.92	2.84	2.20
Japanese	3.40	3.22	0.02	1.44
Russian	7.64	6.93	2.07	1.81
Spanish	9.58	7.89	2.31	2.28

Collecting intensity by subject field or discipline can be examined. Table IV-5 shows the mean number of holding libraries by the 35 subject fields and broad subject groupings defined for the study. Data are only given for the academic library peer groups as the collecting patterns in PG14 so closely parallel those of the ARL libraries, PG1.

In PG1 by broad subject area, the humanities and the reference categories have the same

ALC. Ch. 4, 25

mean number of holding libraries at 6 for non-English language titles. The social sciences have a lower mean number of holding libraries at slightly over 5 libraries per title, while the sciences have over 3 libraries on the average owning a title.

German and Spanish have the highest mean number of holding libraries of all the language groupings in PG1. With 124,034 total titles in the nine-year dataset, the mean number of holding libraries in German is 7.92. Spanish has 111,773 titles with 7.89 holding libraries. Both German and Spanish have a mean of nearly 8 libraries per title. For German the highest means occur in the language and literature grouping at nearly 15 libraries per title. This subject area has the highest mean number of any subject area out of the seven foreign language groupings. The general/ reference category has the highest mean of the broad knowledge groupings. Philosophy, performing arts, general arts and literature all are above the overall mean leading the humanities at 10 libraries per title. The social sciences have nearly 6 libraries per title and the sciences average 4 libraries per title. The collecting is more concentrated in the humanities in German than in either French or Spanish. German has the largest number of titles in the sciences, although not the highest mean number of holding libraries of the seven language groupings, indicating a greater diversity of titles in the sciences in German

In PG1, both French and Spanish have similar means for the four broad knowledge groupings with nearly 9 libraries per title for the general and reference and the humanities groupings. The social sciences have an average of 6 in French and 7 in Spanish. Subject areas in Spanish which average above the mean are bibliography, language and literature and general literature, history of the Americas, and sociology. Political science, business and finance, and the arts in general also are very close to the overall mean and heavily collected.

The collecting in French is similar to the pattern in Spanish, with the highest in the language and literature subject area and philosophy, both over 10 libraries per title. Other areas

above the mean or close to it are the reference and bibliography lines, arts in general, English literature, history (not Americas), and political science. Although German has a much larger number of titles, the mean number of holding libraries in the social sciences and sciences are very close in French and German aside from the heavily collected subjects. The distribution of titles is very similar in both languages while collecting is more broadly spread across subjects in Spanish.

Russian has a lower number of titles than the three major Western language groupings in PG1. The highest mean is in general literature at 9 libraries per title, closely followed by the language and literature grouping at almost 9 per title. History, business and finance and political science are also heavily collected. Russian has the largest number of titles in the sciences of the seven language groupings, although the mean number of libraries is a relatively low 2.71. This is an indication that the collecting in science and technology in Russian is dispersed over a large number of titles. The number of titles for humanities and social sciences are almost identical. The collecting pattern by subject in Russian is similar to that in Spanish: the collecting is more evenly spread over the humanities and social sciences than in French and German in which the titles are more concentrated in a few major subject areas.

Chinese has close to 4 libraries per title overall. The majority of the subject areas have mean numbers of less than 4 with the exception of bibliography, philosophy, languages and literature, history, anthropology, general social science, science, sociology, and geography. The social sciences have the largest number of titles with 35,797, but the humanities have 32,114. Collecting in the sciences averages 2 libraries per title.

Japanese has the lowest mean number of holding libraries of the seven language groupings. It also has the next to lowest number of titles in the sciences. The collecting in Japanese is concentrated in the languages and literature subject area, history (not Americas), and

business and finance. The social sciences have double the number of titles as in the humanities in Japanese with the mean number of holding libraries slightly lower in the social sciences. The highest mean number is in architecture at 5 libraries. Bibliography has close to 5, with all other areas under 4 libraries on the average. Collecting is at a low intensity compared with the European languages. It appears that only a few libraries in PG1 are collecting in Japanese language materials.

The total number of titles and the mean numbers may be so low in the Asian languages as a result of RLG and WLN libraries not being represented in the 1997 CACD database. The absence of a few libraries which collect Asian materials heavily can greatly influence the data for languages which are not collected heavily by the research libraries in the CACD.

While Arabic has the lowest absolute number of titles, the mean number of holding libraries for Arabic averages 5 for all the major knowledge groupings except the sciences at 4 libraries per title. The majority of the titles in Arabic are in the humanities at 17,963; the social sciences have a total of 9,308. The sciences only have 454 titles. The collecting in Arabic is very concentrated. The mean number of holding libraries of 5 per title is higher than the averages for Chinese and Japanese. The libraries collecting in Arabic appear to be acquiring the same titles, resulting in less diversity than in the European language groupings.

While the patterns of collecting do differ by language grouping in PG1, there are commonalities. The highest number of titles in all language groupings is in the language and literature subject areas, not a surprising finding. History is also a heavily collected area as are business and finance, sociology, and political science. The sciences are not heavily collected in any of the language groupings except Russian.

These findings are in agreement with long-standing collection development philosophy and the pattern of scholarship in various disciplines. Research reporting in the sciences is primarily in English. The heaviest use and demand for foreign language materials are in languages and literature, history, political science and area studies. The numbers of titles by language and the mean number of holding libraries reflect this collecting philosophy. There does not appear to be much across-the-board collecting in any language. Spanish, which is the second most spoken language after English in the U.S., is the only language in which the collecting is more broadly dispersed across subject areas. In the other language groupings, the collecting is concentrated in the language, literature, history, politics, and business of the region in which the language is predominant.

Table IV-5

Mean Number of Holding Libraries by Subject, Language, and Peer Group

In PG4, the mean numbers for all foreign language groupings are less than 3 libraries per title for nearly all subject areas. As has been shown in the total number of titles per language by peer group in Table IV-2, the level of collecting in foreign language materials in peer groups 4 and 7 is much lower than in the ARL group. Even in PG1, the lesser collected languages have very low mean numbers. The larger number of libraries in PG4 and PG7 do not result in a higher mean number of holding libraries in the foreign languages as they do in English language materials. When the number of libraries in these two peer groups are taken into consideration, the level of foreign language collecting is extremely low.

The mean numbers for all non-English as a category in PG4 are highest for general/reference and the humanities. Only the history of the Americas has above 2 libraries per title in the social sciences. The highest mean number of all subjects is in mathematics with 5 libraries per title. The only other subject area with above 3 libraries per title is the languages and literature category. By language grouping, the collecting patterns in PG 4 are very similar to those in PG1, the ARL libraries. German has the largest number of titles, followed by Spanish, French, Russian, Chinese, Japanese, and Arabic. The humanities average over 3 libraries per title in both French and German which have close to the same number of titles in languages and literature. There are a considerable number of titles in philosophy, religion, and the performing arts in German. In both German and French, history (not Americas) has by far the largest number of titles in the social sciences. There are no real collecting emphases in the other social sciences in either language. The social sciences have half as many titles in French as the humanities, and nearly 40% fewer titles in German. The mean number of holding libraries in both languages in the social sciences is slightly less than 2 libraries per title. The number of titles in the sciences is highest in German at 2,167. The overall averages in the sciences are 2.84 holding libraries per title for German and 2.70 for French.

The pattern for Spanish in PG4 is similar in that there are only two thirds as many titles in the social sciences as in the humanities. The concentration in the humanities is in the languages and literature area almost exclusively. In the social sciences business has comparatively more collecting intensity in Spanish than in French and German. All subject areas in the social sciences have less than 2 holding libraries per title. The number of titles in the sciences in Spanish is almost exactly the same as in French. The overall mean number of holding libraries is slightly lower than French and German at 2.31 libraries per title.

Russian has barely 2 libraries per title overall in PG4. Only 37% of the titles are in the social sciences and the mean number of holding libraries is highest in the languages and literature area and in the humanities. Unlike in PG1, there are a very low number of titles in the sciences in PG4 with the exception of a mean number of nearly 7 libraries per title on a small number of titles in mathematics.

The mean numbers by language grouping and subject for PG7 are shown in Table IV-5. The collecting patterns in PG7 are very similar to those of PG4, but with lower mean numbers. PG7 has almost exactly half the number of non-English language titles as PG4 – 90,509 versus 179,459. The number of titles in Japanese, Russian, and Spanish are close to being identical in both PG4 and PG7. But PG4, has almost twice the number of titles in German as PG7. Arabic is virtually not collected at all in PG7, with only 272 titles to over 5,000 titles in PG4.

While the number of titles is much larger in PG4, the mean numbers are very similar with the exception of German, which is much more heavily collected in PG4 than in PG7.

The collecting patterns for both non-ARL peer groups 4 and 7 indicate that a very small number of libraries within the peer groups are collecting foreign language materials. If we look at Table IV-4, it seems that these libraries may be collecting the same titles as the ARL libraries in that the mean numbers for PG14 appear to be a total of the PG1 means and the mean levels of

PG4 and PG7. The two non-ARL peer groups seem to be adding to the mean numbers for the ARL libraries with respect to foreign language holdings.

The mean number of holding libraries is an indication of collection overlap – how many libraries on the average are adding the same titles. The next section analyzes the foreign language groupings by an opposite measure, the number of unique titles.

Unique Titles by Language: English, Non-English

It is possibly more important with lesser used materials to be concerned about the diversity of available titles than that titles be widely held. We have seen that the mean number of holding libraries for English language materials is three to one to non-English language materials. While there can be cause for concern if the mean numbers are low, there can also be reason for concern if the mean numbers indicate “collection convergence” -- that is, more libraries buying the same titles and fewer distinct titles being acquired. The number of unique titles is an indication of diversity and is especially important for foreign language materials since they are less used, but also less collected.

As already defined in chapter III, a title is unique within a peer group if only one library within the group owns that title. A title may be unique in more than one peer group, that is, one library in each peer group could own the same title, but it would be unique in each peer group because only one library in each group owned the title. The number of unique titles in Peer Group 14, the CACD, database means that only one library of all of the libraries in the database owns those titles.

The number of unique titles for each peer group by year are shown in Table IV-6. The total number of unique titles is divided into two categories of English and non-English language titles. The percentage of the total number of titles for each year that are unique are shown in the

“Percent of Total” column.

From Table IV-6, it can be seen that all unique titles, that is all languages, comprise 23.5% of the CACD peer group 14 by publication year in 1987. Slightly more than half of all unique titles are foreign (non-English) language titles. Proportionately, for PG14 non-English language unique titles make up 3% of all titles, that is, 3% of the 1997 CACD database is comprised of foreign language titles for which there is only one library location symbol.

For the ARL group (PG1) the percentage of total in foreign language unique titles was near 13.5% in 1987. For the ARL group the number of unique titles increases every year until 1992, at which point the total number of titles decreases. The percentage of unique titles fluctuates in an increasing trend from 1987 to 1995. Both the number of non-English language unique titles increases and the percentage share of all titles that are unique increases. This would seem to indicate that additional holding symbols are accumulated over a long time span for foreign language titles as there are more unique titles each year going forward in time. In other words, there are possibly more unique titles each year closer to the production of the CACD database because other libraries have not added their location symbols to the records, although cataloging lag would seem to be most operative in the addition of the cataloging to the database rather than the addition of library holdings.

The percentage of total comprised of unique titles does not differ widely between the divisions of English and non-English in PG1. That is, the percentage of English language titles which are unique ranges from 23-25%. The range for non-English titles is from 27-30%. There is a difference of 4-5 percentage points from the percent of total between the two divisions. While non-English titles have a higher percentage of total in unique titles, English language titles average close to one-fourth in unique records. That is, of all titles in English, approximately one in four is owned by only one library in PG1. In terms of resource sharing, this could be regarded

as a high level of collection convergence in that 75% of titles are owned by more than one library, but the percentage of unique titles--diversity--is rather low.

Table IV-6

Unique Titles as Percentage of All Titles: English, Non-English

In PG14, there is considerable difference in the absolute number of English language titles between 1987 and the last years in the study. The number of English language titles increases by 19% between 1987 and 1992. It can be seen that the ARL libraries have close to the same percentages of unique titles by language divisions as does the database. Again, analyzing the data for PG1 suffices for PG14 as well.

In PG4, the ratio of English to non-English unique titles is predominately English. The pattern of rise and fall in the absolute numbers is similar in both categories, rising from 1989-1993 and then decreasing.

The absolute numbers of unique English language titles for PG7 are very close to those for PG4. But the number of unique non-English language titles for PG7 is far below that of PG 4. PG7 has approximately 40 % fewer unique foreign language titles than PG4. The percentage of non-English That is?

Because the measure of unique titles to total is subject to cataloging lag, especially with foreign language titles, the chief value in this measurement is in comparing patterns among peer groups. If we look at the percentage of English language titles that are unique in the three academic library peer groups (Table IV-6) we can see that in all three the percentage of total is in the 23-26% range. The three peer groups have the same pattern in percentage of unique English language titles. But the situation is remarkably different with respect to non-English language titles. In both peer groups 4 and 7, the percentage of total for unique non-English titles is more than double that of English language titles. The percentage of unique non-English language titles fluctuates considerably from year to year in PG4. It is relatively steady within a range of 63-66% in PG7. Thus, the three academic library aggregated collections are very similar with respect to collecting patterns for English language materials. But the two non-ARL library peer groups are very different from the collecting pattern for foreign language materials in the ARL group. Not

only are the percentages of total for each division in peer groups 4 and 7 similar, the absolute numbers of unique English language titles are on the average, only 5,000 titles apart. But the numbers for unique non-English language titles are much more divergent in those two peer groups -- 10,000 titles apart. Thus while the collections are predominantly English language materials, it is the high percentage of non-English unique titles which causes the higher overall percentage of total for unique titles in the two non-ARL peer groups.

The individual language groupings are analyzed next for percentage of unique titles.

Unique Titles by Foreign Language Groupings

The total number of titles, number of unique titles, and percentage of total that are unique are displayed in Table IV-7, for the seven foreign language groupings in the three academic library peer groups by subject areas. Beginning with PG1 we can see that the percentage of total that is unique varies considerably by language grouping. Japanese has the highest overall percentage unique followed by Chinese. French, German, and Russian which have close to the same percentage unique at 25-26 percent. Spanish is slightly lower with 22% and Arabic has the lowest percentage unique at 18 percent.

There is also variance in the language groupings in the percentage unique by broad knowledge groupings. Japanese, with the highest overall percentage of unique titles of the seven language groupings, has the highest percentage of unique in the sciences at 54 percent. The social sciences have the next highest percentage unique at 41% followed by the humanities at 35% and the general category at 32%. If we compare this measurement with the mean number of libraries data we see that the percentage unique is highest in the areas in which the mean number of libraries is lowest. Because Japanese is the next to lowest in collecting intensity, the absolute number totals are not large, making the number of unique titles very small except for the

aggregated data in the broad knowledge groupings.

Although there is a difference of 16,000 in total number of titles between Chinese and French, in PG1 the number of unique titles for both languages is virtually the same at 23,000. Since Chinese has the smaller number of titles, the percentage unique is larger at 31% than French which has 26% unique. Both languages have the same percentage unique for the reference category, but the percentages vary for the three large broad knowledge groupings. Chinese has 30% unique in the humanities while it is much lower at 24% for the humanities in French. Chinese has the same percentage in the social sciences as the humanities at 30% while in French the percentage of unique titles in the social sciences is 26 percent. For the sciences Chinese has 53% unique, while for French the percentage is 37%. In both languages the humanities and social sciences have close to the same percentage within the language grouping. In both languages the sciences have a much higher percentage of unique titles.

German has the largest absolute number of titles in PG1, but the percentage of unique titles is the same as for French and Russian. As in the others, the highest percentage unique is in the sciences. One in five titles in the humanities in German is unique, lower than the 31%, or almost one in three titles in the social sciences. The humanities and social sciences have only 4,000 titles difference in the absolute total and approximately 5,000 titles difference in number of unique titles in the humanities. But the difference in percentage share of unique titles is 10 percentage points less in the humanities which has the larger number of titles.

Arabic, French, and German have more titles in the humanities than the social sciences. Russian has an equal number in both. The situation is reversed for the remaining language groups, Chinese, Japanese and Spanish. In Arabic and French, the percentage unique within the language grouping for the humanities and social sciences is the same, but there is a 10 percentage point difference between the humanities and social sciences in unique titles in German.

Arabic, Chinese, French, Russian, and Spanish all have very similar percentages of unique within each language groupings for both the humanities and the social sciences. In Japanese and German, the difference in percentage of total between the humanities and social sciences is much greater. The humanities have the lower percentage of total in both. In German, the social sciences have 10 percentage points more than the humanities in unique while in Japanese the differential is five percentage points.

From these data it appears that the languages most heavily collected have the lowest percentages of unique titles. It appears that the same titles are being purchased, thus higher mean numbers in the more heavily collected languages and fewer unique titles. The percentage unique in foreign language titles for PG14 in effect remove the overlap between the three academic library peer groups. The percentage of unique titles in PG14 is virtually identical to the percentage unique in PG1 showing that the unique titles in the non-ARL peer groups are only unique in those peer groups, but appear to be titles which are also owned in the ARL group. Thus, the non-ARL libraries do not appear to be adding unique foreign language titles to the information commons of the aggregated collections.

Table IV-7

Unique Titles by Foreign Language Grouping

The pattern of percentage of unique titles varies within the different language groupings. They fit with the mean number of holding library data in that the percentage of unique titles is highest in those areas with a lower mean number of holding libraries. The patterns for collecting by subject are similar in all of the language groupings in that the concentrations are in the same subject areas.

Data are shown for the percentage of unique titles by subject in both peer groups 4 and 7 in Table IV-7. It follows that with the low intensity of collecting in foreign language monographs in both peer groups that the percentage of total comprised of unique titles is extremely high. In PG4, only French (42%), German (47%) Russian (58%), and Spanish (55%) have overall percentages of unique titles below 80 percent. In PG7, Arabic and Chinese have over 90% in unique titles. German is the third highest at 82% with only French (60%), Russian (67%), and Spanish (63%) below 80 percent.

In PG4, all of the languages have a larger total number of titles in the humanities than in the social sciences. The percentages of unique titles in the broad subject groupings do vary by language but, in all but Arabic, the percentage of total in unique titles is lower in the humanities than in the social sciences and sciences. In Arabic there is only one subject area which has below 80% in unique titles and that is the history of the Americas. Chinese has only a few subject areas with below 80% unique: bibliography, religions, visual arts, architecture, geography, general social science, and library and information science. In French, the majority of the subject areas are in the 30% to 60% range, with the exception of English literature, recreation, business and finance, and the majority of the sci/tech areas, which are all over 60% of total in unique titles. In French and German, the humanities subjects have among the lowest percentage in unique titles, reflecting stronger collecting intensity. The range is the same in German with only American literature in the humanities having over 60% in unique titles. In German the social sciences all

have over 60% in unique titles with the exception of psychology and history (not Americas). The sci/tech areas all have over 60% except computer science and general science. The majority of the subject areas in Japanese have over 80% unique with the exception of architecture and philosophy.

The patterns in Russian in PG4 are similar to the other languages in that there are only three subject areas with below 60%. One of these is the language and literature subject line which has a large number of titles. At 47% percent in unique titles this subject area is responsible for the humanities in Russian having a relatively low 51% in percentage of unique titles. Spanish follows the same pattern as Russian in that the language and literature areas have a much lower percentage of unique titles.

In PG7, French is the only language to have less than 60% overall percentage of total in unique titles. German has 61%, Spanish 63% and Russian 67 percent. Arabic and Chinese have over 90% with Japanese at 82 percent.

In PG7, the majority of titles in non-English are in the humanities which do have a lower percentage of unique titles than the sciences and social sciences. Spanish has a much larger proportion of total titles in the social sciences, but the percentage of unique titles in the social sciences overall is nearly 75 percent. As the only appreciable collecting in PG7 is in French, German, and Spanish, the collecting patterns for the peer group reflect the relative absence of area studies with the possible exception of Latin American which would account for the large amount of material in the social sciences in Spanish. The collecting in French and German is heaviest in the core humanities areas of religion, the arts, and languages and literature and in the social sciences in history. Foreign language collecting in PG7 is in mainstream scholarship, rather than professional areas and social sciences and science/technology. Spanish is mostly heavily collected followed by German and then French. The high percentage of unique titles and

the low mean number of holding libraries indicate that out of the 227 libraries in PG7, only a small number of libraries are collecting foreign language materials to any extent. The subject areas with the highest number of titles and lower percentage of unique titles are areas in which the collecting is more widespread among the libraries and there are titles which the libraries are buying in common. But for the majority of the languages and subject areas, the collecting is scattered among a few titles and even fewer libraries.

Summary

The data for non-English imprints exhibit some characteristics and trends which are similar to those of the English language imprints. The collecting patterns of the ARL libraries, PG1, and those of the 1997 CACD database, PG14, are similar. The much higher level of collecting intensity in foreign language monographs of the ARL libraries seems to exert even more influence upon the database with respect to foreign language collecting patterns than English language collecting patterns.. The very high levels of English only collecting in the non-ARL peer groups also exert an influence upon the database such that the percentage of English language titles averages 57% in the 1990s in PG14 versus an average near 48% in PG1.

The ARL libraries actually maintain the most consistent English/non-English ratios of the four peer groups. PG4 exhibits the most pronounced downtrend in non-English imprints. In all peer groups the highest percentage of share for non-English imprints occurs in 1988, and with few exceptions the percentage of English language titles increases beginning in the 1990s. The differences are slight, less than one percentage point decrease in non-English titles in many years and peer groups. English language titles have three times as many holding libraries as non-English on the average in PG1.

Within the seven foreign language groupings, collecting activity shown in PG14 is lower

from 1987-1991 than in the later years of 1992-1993. Not surprisingly, German, Spanish and French are the most heavily collected languages in that order. French has the most consistent total, almost virtually the same in each year until 1995. Spanish has the largest increase in total number of titles at nearly 3,000 more titles in 1992 than 1987. While the total number of titles fluctuate within the language groupings by year, the percentage share of total for each language remains relatively consistent from year to year across all peer groups.

By broad knowledge grouping, the ratio of English to non-English follows a similar pattern in the ARL and database groupings, and a very different pattern in the two non-ARL academic library peer groups. In PG14 and in PG1, the ratio in the humanities has a higher percentage in non-English than in English. The preponderance of English language collecting in PG4 and PG7 result in predominantly English language imprints in all the major knowledge groupings for those peer groups.

In the social sciences in peer groups 1 and 14, English holds a slight advantage over non-English in a 51/49 or 52/48 ratio in most years. In PG4 the ratio in the social sciences is close to 80/20 and it is 90/10 for PG7. The sciences have the highest increase in English language materials of the broad knowledge groups in both PG1 and PG14, from 80/20 in 1987 to near 85/15 in 1993. The level of foreign language collecting in the sciences in PG4 and PG7 is so low that the science collections in those peer groups are, for all practical purposes, totally in English.

The mean number of holding libraries show the levels of intensity in foreign language collecting by peer group. Although the number of libraries in each peer group varies greatly, the mean number of holding libraries per title in the foreign language groupings are very similar in peer groups 1 and 14 and in peer group 4 and 7. It is obvious from the mean number data in Table IV-4 that the ARL library collections hold the preponderance of non-English language titles and set the patterns for mean number of holding libraries in the database, PG14. The other

two peer groups, 4 and 7, have extremely low levels of foreign language collecting. In spite of the different number of libraries in those two peer groups, the mean numbers by language grouping in the two peer groups are very similar, meaning that collecting intensity is much lower in the medium-sized library group.

The present study has found that collecting of foreign language titles is very low in the large academic library group, PG4, and almost non-existent in the medium-sized libraries in PG7. Schreiner-Robles^{ix} conducted a survey in 1986 of collecting methods for French, German, and Spanish literatures at twenty-nine medium-sized academic libraries. The level of degrees offered in both English/American and the foreign literatures were master's and bachelor's. Monographic holdings of the libraries ranged from 240,000 to 750,000 volumes, similar in size to the libraries in PG7 in the CACD. Findings were that "On average, ten libraries order 30 to 50 percent fewer book titles for all three degree programs in French, German, and Spanish than for one degree program in English."^x Curriculum was ranked as the most important factor in making collection decisions and faculty were heavily depended upon for purchasing original language books, with only 4% of the librarians using foreign language review tools. Although there were questions in the survey about the number of acquisitions by language and translation, no data were given in the study report on the ratio of English translation to original language titles acquired. Only one library reported using foreign book dealers to supply original language materials. The results of the survey indicate that medium-sized academic libraries rely on domestic approval plans and U.S. reviewing sources as collection development methods for foreign literatures.^{xi} Given that none of the institutions offered Ph.D degrees in either English/American or the foreign literatures, these may be appropriate collecting methods. These findings describe the collection methods in medium-sized libraries and assist in explaining the low collecting intensity in foreign languages found in the present study.

For the period 1987-1993, foreign language collecting did not decline, but rather increased in terms of total number of titles per year. The percentages of total, both for the individual language groupings and the aggregate grouping of non-English, did not change significantly. This is due to the overall total number of titles increasing in all peer groups through 1993, that is, the number of English language titles increased as well as the foreign language titles so that the percentage shares between the two remained relatively consistent. The collecting patterns by numbers show that a recovery seemed to be taking place from the down years of the latter 1980s.

Notes for Chapter IV

i. Mark L. Grover, "Cooperative cataloging of Latin American books: the Unfulfilled Promise," *Library Resources & Technical Services* 35 (1991): 406-15.

ii. Cecilia S. Sercan, "Where has All the Copy Gone? Latin American Imprints in the RLIN Database," *Library Resources & Technical Services* 38 (1994):561-69.

iii. Konstantin Gurevich, "Russian Monographic Records in the OCLC database: A Crisis in Shared Cataloging," *Library Resources & Technical Services* 35 (1991): 459-61.

iv. Jai-hsya Tsao, "The Quality and Timeliness of Chinese and Japanese Monographic Records in the RLIN database," *Library Resources & Technical Services* 38 (1994): 60-63.

v. Viveca S. Seymour, "A Survey of East European Monographic Records in the OCLC Database," *Library Resources & Technical Services* 38 (1994): 275-79.

vi. "The Crisis in Libraries' Collection Coverage of Global Publishing: An Update," ARL Newsletter, special issue on Global Resources, no. 206 (October 1999): 8-11.

2. Anna H. Perrault, "The Shrinking National Collection: a Study of the Effects of the Diversion of Funds from Monographs to Serials on the Monograph Collections of Academic Libraries," *Library Acquisitions & Theory* 18, no.1(1994).

viii. Ibid.

ix. Rebecca Schreiner-Robles, "Collection Development in Foreign Literatures at Medium-sized Academic Libraries," *LRTS* 32, no.2 (April 1988): 139-147.

x. *Ibid.*, 142.

xi. *Ibid.*, p.