March 2021

Mapping the Intellectual Structure of Linguistics Research Over 101 Years (1919–2020)

Fasih Ahmed
COMSATS University, Islamabad, Pakistan, ahmadfasih33@gmail.com

Murtaza Ashiq
Islamabad Model College for Boys H-9, Islamabad (Pakistan), gmurtazaashiq00@gmail.com

Shafiq Ur Rehman
Imam Abdulrehman Bin Faisal University, P.O.Box 1982, Dammam, Saudi Arabia, surehman@iau.edu.sa

Shakil Ahmad
Prince Sultan University, Riyadh, Saudi Arabia, shakil@psu.edu.sa

Shaista Rashid
Prince Sultan University, srashid@psu.edu.sa

Mapping the Intellectual Structure of Linguistics Research Over 101 Years (1919–2020)

Fasih Ahmed
Department of Humanities
COMSATS University, Islamabad, Pakistan.
Email: ahmadfasih33@gmail.com
https://orcid.org/0000-0002-9357-7073

Murtaza Ashiq
Lecturer (Library & Information Science)
Islamabad Model College for Boys
H-9, Islamabad (Pakistan).
Email: gmurtazaashiq00@gmail.com
https://orcid.org/0000-0002-3926-0673

Shafiq Ur Rehman
Associate Professor, Deanship of Library Affairs,
Imam Abdulrehman Bin Faisal University,
P.O.Box 1982, Dammam, Saudi Arabia.
Email: suRehman@iau.edu.sa
Contact no: +966 56 535 3240
ORCID ID: 0000-0002-8169-0132

Shakil Ahmad
Research Services Librarian
Central Library, Prince Sultan University
Riyadh, Kingdom of Saudi Arabia
Email: shakil@psu.edu.sa
ORCID ID: 0000-0001-7775-4917

Shaista Rashid
College of Humanities
Prince Sultan University
Riyadh, Kingdom of Saudi Arabia
Email: srashid@psu.edu.sa

Abstract

This article reviews 101 years (1919–2020) of research in the field of linguistics. The main objective of this paper was to analyze the published research and citations on linguistics and present the results from the perspective of growing trends, most productive and influential countries, organizations, authors, research journals, keywords, authorship patterns, and international collaborations. The subject category of linguistics selected in the Web of Science database, for this purpose, 6144 publications have been retrieved from the Web of Science database. This retrieved data is analyzed using visualization tools and dividing it into two phases on a temporal basis. Overall, the first phase, from 1919-1999, shows a lower number of research publications in linguistics, whereas the second phase (2000-2020) with 5585 publications and 41195 citations shows a significant growth in linguistics research. Among them, the highest number of publications is from the US that is followed by the UK. Noticeably, the consideration of the authorship of these researches indicates a strong preference for single authorship (with 3803 articles) in comparison to co-authorship (with only 1463 articles). On the basis of frequency-based keyword analysis, it is observed that the interaction of linguistics with other field s such as computer science, psychology, and the law has been particularly productive for the establishment of new subfields in linguistics which include computational linguistics, cognitive linguistics, and forensic linguistics.

Keywords: Linguistics; language; bibliometric analysis; research productivity

1.   Introduction

Linguistics is the scientific study of language. Linguists apply the scientific method to conduct formal studies of speech sounds, grammatical structures, and meaning across the world's 6,000+ languages (America, 2020). Linguistics covers both written and spoken aspects of language for research. It is a multifaceted subject covering sociolinguistics, language theory, language history, phonetics, semantics, and rhetoric (Reference, 2020). The contemporary issues in linguistics focus on public policy, students' issues, endangered languages, ethics, human rights, women and gender, ethnic diversity, higher education, K-12 education, and public outreach. Working on these issues, linguistics has come across multiple other fields that resulted in the subcategories of linguistics (America, 2020). Linguistics as the scientific inquiry of languages covers various aspects of research that involve language acquisition, language learning, language revitalization, language extinction, the orthography of languages, language policies, etc.

The effect of globalization has greatly modeled the way linguistics prevails in the present times, from simple language learning to the process of natural language processing where machines are able to communicate. Linguistics works in two paradigms can be regarded as a core paradigm and intersecting paradigm. The core paradigm linguistics deals with its core areas such as phonetics, phonology, syntax, semantics, typology, discourse analysis, pragmatics, etc. The intersecting paradigm intersects with many other disciplines, such as psycholinguistics, sociology as sociolinguistics, history as historical linguistics, computers as computational linguistics, and neuroscience neurolinguistics and ecology as ecolinguistics, etc. This position of linguistics makes it stronger to deals with its core areas as well as with other disciplines in research (O'Grady et al., 1997). From this point of view, publications in linguistics have seen a dramatic change from the first periodical in linguistics (i.e., *The Modern Language Journal*, 1916) to the present time (Mohsen et al., 2017).

The present study aims to observe the modern trends in the fields of linguistics since little is known about the existing research trends in linguistics which much needed area of inquiry specifically for the meaningful research productivity. Existing studies in linguistics have mainly focused on the situation of languages where some languages are on the verge of extinction, and linguists across the world are focusing on revitalizing these languages (Liljegren, 2018). Among them , local languages in Asia and Africa prominently, observed as losing their speakers at a rapid speed (Weinreich, 2010).

Similar to this trend various bibliometric studies in linguistics have been conducted. Arık (2015) has conducted a bibliometric analysis of research works conducted in the field of linguistics, represented in WoS (Web of Science) over a period of 1900 t0 2013. Arık (2015) Bibliometric analysis reflects overall progress in linguistics from 0.8% to 4.52%. The bibliometric analysis of Ching (2012) using ISI (Institute of Scientific Information) Web of Science only considers language studies conducted in Taiwan over a time period of 40 years. This study shows a surge in language research but ignores various aspects of linguistics as it only centers on language studies especially in learning and teaching. Jaber et al. (2017)'s bibliometric analysis of Scopus published articles of English linguistics of Arab scholars and presents research productivity in Arab countries. Lei and Liao (2017)'s bibliometric analysis of linguistic productivity in China, Hong Kong, Taiwan, and Macau reveals that Hong Kong is the leading region in linguistic research conducted in these four regions from 2003 to 2010. Then China was noticed as the leading country in research in linguistics. Their analysis presents the research productivity of only four countries but the consideration of a wider canvas of linguistic research is still missing. Another study reports the bibliometric analysis of linguistic publications published in Web of Science from 2005 to 2014, with the USA as the highest number of publications in research in linguistics. Ezema and Asogwa

(2014)'s bibliometric analysis focuses on sources cited in articles published in two journals of linguistics from 2001 to 2010. One of the important aspects of their findings reveals that the degree of collaboration in linguistics was noticed very low. However, the finding was limited only to the two selected journals and cannot be generalized to the overall progress of linguistics. The bibliometric analysis of *system* journal in applied linguistics has been conducted by Lei and Liu (2019a). Their findings report publications with a focus on language learning and teaching practice issues and the use of technology as the prominent topics over a time period of four decades. Their bibliometric analysis presents the state of research with respect to one journal only that too mainly focuses on issues in applied linguistics.

It is argued that the above bibliometric studies conducted in linguistics are limited in scope on three grounds. Firstly, they only reflect an increase in citation index in linguistics but does not consider the emerging fields in linguistics such as computational linguistics and natural language processing, who plays an important role in meaning making in the field of technology surveillance. Secondly, most of the bibliometric studies are limited to the research productivity of linguistics with respect to specific countries. It is vital to look at overall progress in linguistics beyond the traditional borderlines. Thirdly, some of the bibliometric studies in linguistics are limited in the sense of being conducted by focusing only on a specific number of journals related to linguistics. To the best of the authors' knowledge, no single study is available that deals with the state of research in linguistics over an extended time period, as in the case of the present study. Therefore, to addresses all these limitations of the previous researches mentioned above, present bibliometric study of linguistics beyond the scope of a specific country or specific journal is a much needed enterprise.

Bibliometric analysis is vital to determine contemporary trends in a research discipline, which helps in mapping the overall shape of a research discipline. This mapping out brings forth evaluation and comparison of research directions, which is highly useful for research to determine the overall image of a research discipline (Ching, 2012; Korevaar, 1996). In other sense, the bibliometric analysis leads to a specific research profile of a discipline in a country (Clarke et al., 2007; Mohsen, Fu, & Ho, 2017). This profile highlights the research performance on the basis of institutions, research journals and other related functions to monitor research performance. Citation coverage, spanning over a specific time period, also helps in analyzing the research ranking of institutions, journals and the contribution of research scholars from a specific country in the overall research articles cited in that specific field (Ezema & Asogwa, 2014; Lei & Liao, 2017). Citation coverage apart from the ranking of institutions also proves helpful in job hunting for new graduates, promotion of faculty members and respectability of researchers (Arık, 2015; Ho, 2014). This bibliometric analysis focused on the following research questions.

1. What publishing trend has been dominant in the field of linguistics?
2. What are the preferred journals of researchers in linguistics?
3. What are the most productive authors, institutes, and topics in linguistics?
4. What are the authorship patterns of researches in linguistics?
5. What are the frequently used keywords in linguistics research?

2. Methodology

Bibliometric method is chosen to identify the prevalent trends in existing linguistic research articles. Bibliometric analysis is a quantitative method to analyze the research productivity and citations of published academic work. The researchers used Web of Science core collection (WOS) database to download bibliographic data in WoS subject category "linguistics". It is worth

mentioning that WOS is the most authentic, reliable, and accurate indexing and abstracting database of peer-reviewed literature. The data was retrieved from WoS on May 22, 2020, at Imam Abdul Rehman Bin Faisal University. A total of 6144 papers were retrieved, including articles (4693), Review articles (211), early Access articles (45), proceedings Paper (1164), Book and Book Chapter (34), and editorial Material (2). Moreover, reports, news items and correspondence were excluded from the study.

The data files were exported, and analysis was performed using data visualization softwares (VOS viewer, biblioshiny, HistCite, ScientoPy) and MS Excel spreadsheet. A few fields/abbreviations are used in the various columns. Thus, TP stands for total publications, TC describes the total citation whereas CI stand for citation impact which is the average number of citations that a specific publication has received. The citation impact was calculated by dividing the total number of citations by the total number of publications.

3. Data Analysis

3.1 Overview of Influential Aspects of the Linguistic Literature

The most influential aspects of the Linguistics literature have been identified such as publishing trends (Figure 1), top countries (Table 1, Figure 2-3), institutions (Table 2), authors (Table 3), journals (Table 4), authorship pattern (Figure 4), keyword analysis (Figure 5 & 6) and three-factor analysis of major aspects on linguistics literature (Figure 7 & 8). The initial analysis showed that 6144 records in the WOS dataset were published by 496 journals, written by 7982 authors, affiliated with 199 institutions and 114 countries. These documents received 49178 citations published in 21 languages.

3.2 Analysis of the Overall Growth Trend

Figure 1 shows the year-wise frequency of publications and citations published from 1919 to 2020. The 100 years of linguistic publications and citations were divided into two phases. In the first phase (1919-1999), there were 559 publications, 8083 citations. The most productive year in that era was 1999, when 39 publications appeared. The best year with regards to citation was 1996, which received 1619 citations. The citation impact of the first phase was 14. The second phase (2000-2020) produced 5585 publications and 41195 citations. The significant growth (N= 2640) has been observed from 2015-2019. The years 2019 and 2018 were marvelous as in those year's total, 1141 research publications were produced. The year 2019 is excellent as in that year, 571 publications were produced. The year 2010 was the most cited year when 4104 citations were received.
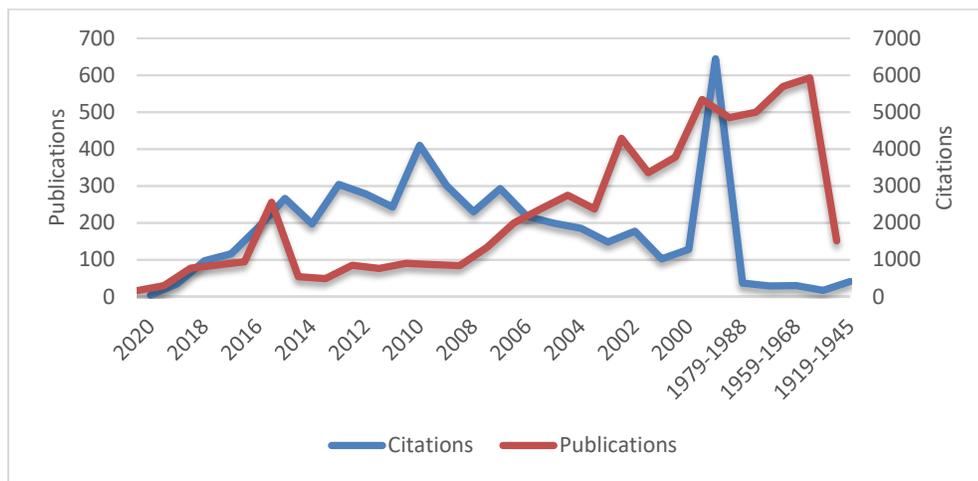
Figure 1: Publication and citations trend on Linguistics Research (1919-2020)

3.3   *Influential Countries*

Top 10 countries producing publications on Linguistic are presented in Table 1. Only two countries that produced over 600 publications are United States of America and United Kingdom. United States of America is on the top of the list and far ahead in top 10 most productive countries producing 1084 publications, 15660 citations, and 14.36 citation impacts. United States and United

Kingdom also remain the top two countries for producing maximum publications from 1995 to 2020 whereas, Spain is replaced by China producing higher citations in the last four years from 2016 to 2019 (Figure 2).

Table 1: Top 10 influential countries on linguistics research

Sr. No.	Country	TP	Percent	TC	Citation Impact
1	United States of America (USA)	1084	17.6	15660	14.361
2	United Kingdom (UK)	664	10.8	9812	14.326
3	China	391	6.4	2133	4.517
4	Spain	372	6.1	1532	4.272
5	Germany	350	5.7	2217	6.633
6	Australia	302	4.9	2522	9.049
7	France	241	3.9	1186	5.829
8	Canada	205	3.3	3281	13.698
9	Netherlands	155	2.5	1377	8.376
10	Belgium	133	2.2	1174	8.552

In addition, Belgium (bottom of table 1) also replaces Italy by producing higher publications (Figure 2). United Kingdom ranks second on the most productive countries list by producing 664 publications, followed by China with 391 publications, Spain with 372 publications, and Germany with 350 publications. It is noted that though Canada is at number 8 in the list with 205 publications; however, the impact of its publications is 13.69. USA and UK have maximum citation impact.
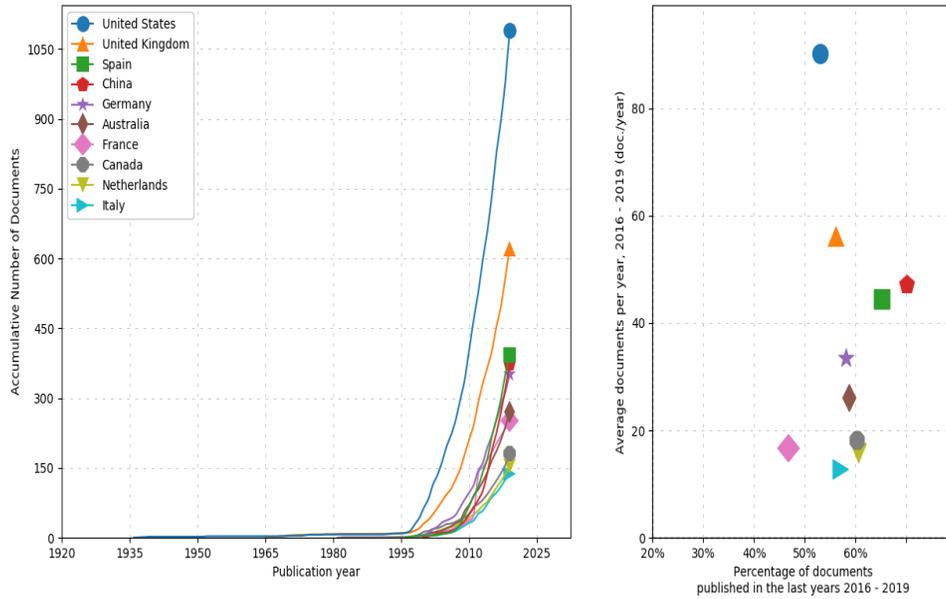
Figure 2: Top ten countries under different chronological period

3.4   *Country Collaboration Map on Linguistic Literature*

Figure 3 shows the country collaboration map on linguistic literature globally. There are six collaborator countries that have over 20 mutual publications, whereby US has privilege to collaborate with five of those countries. Both US and the UK emerged as top collaborator countries mutually producing 47 publications, followed by Germany and USA with 34 publications, USA and Canada with 32 publications, USA and China with 25 publications, Germany and UK with 23 publications, and Netherlands and USA with 21 publications.
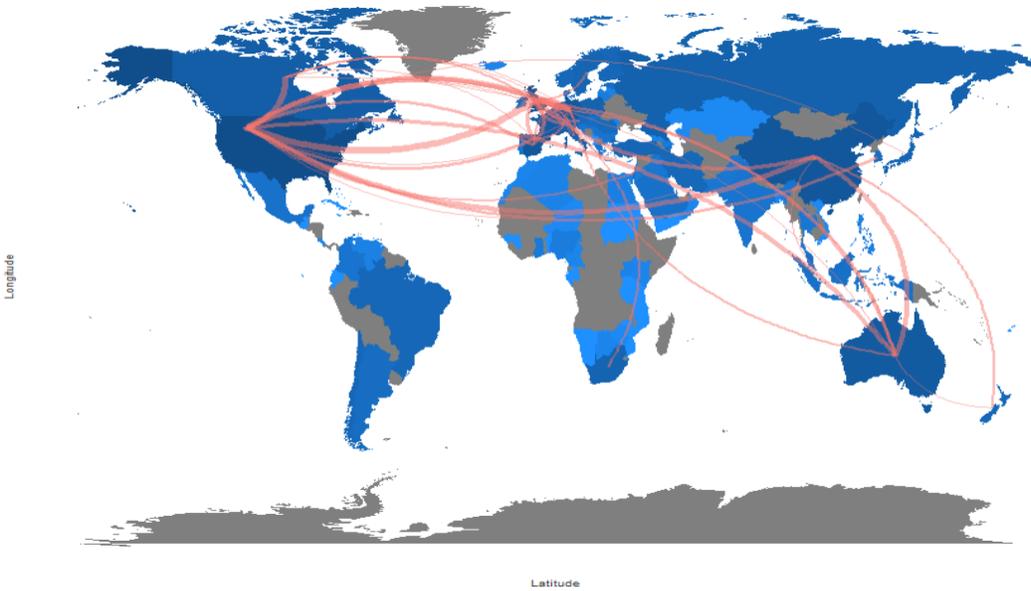
Figure 3: Global country collaboration map on linguistic literature

3.5 Highly Influential Organizations

Top 10 highly influential organizations in linguistics are given in Table 2. Three outstanding institutions have over 50 publications out of which University of Michigan, USA, is considerably top of the list with notable 59 publications, 1739 citations of those publications and overall 29.47 citation impact, followed by The French National Centre for Scientific Research (CNRS), France with 56 publications and 285 citations, The University of Sydney, Australia, with 52 publications and other organizations as shown in Table 2. Furthermore, it is noteworthy that out of these ten organizations in the world, three institutions belong to the U.K, two–to Australia, Hong Kong. Furthermore, The University of Edinburgh, UK and University of British Columbia, Canada rank 4 and 9 in that list with 56 and 39 publications, but it is distinguishable that its publications have the highest citation impact, i.e., 40 and 31.56, respectively.

Table 2: Top Ten Highly Productive Organizations

Name of Institution and Country	TP	TC	Citation Impact
University of Michigan, USA	59	1739	29.47

The French National Centre for Scientific Research (CNRS), France	56	285	5.08
The University of Sydney, Australia	52	506	9.7
The University of Edinburgh, UK	49	1960	40
University of Birmingham, UK	45	471	10.46
The University of Hong Kong, Hong Kong	45	431	9.57
Lancaster University, UK	44	592	13.45
University of Helsinki, Finland	41	308	7.51
The Hong Kong Polytechnic University, Hong Kong	39	238	6.10
Macquarie University, Australia	39	280	7.17
The University of British Columbia, Canada	39	1231	31.56
Ghent University, Belgium	39	279	7.15

3.6 Most Prolific Authors

The most prolific author's data revealed that there are six authors that have 10 or over 10 publications (table 2). The list of most prolific authors shows that Gries ST is the most productive author with 16 publications, 499 total citations, and 11 h index. Ellis NC, Geeraerts D., Liu HT, and Newmeyer FJ have produced the same number of 11 individual publications however, Ellis NC publications have comparatively highest citations. The four authors at the bottom of top-ten-authors have produced 9 publications individually.

Table 3: Authors Impact

| Author	TP	TC	FY of Publication	H_index	G_index	M_index
Gries ST	16	499	2005	11	16	0.6875
Ellis NC	11	1069	2006	10	11	0.666667
Geeraerts D	11	180	2003	5	11	0.277778
Liu HT	11	117	2010	5	10	0.454545
Newmeyer FJ	11	337	1986	6	11	0.171429
Joseph JE	10	41	1996	4	5	0.16
Albury NJ	9	30	2015	3	5	0.5
Crossley SA	9	271	2009	9	9	0.75
Martin JR	9	139	2013	6	9	0.75
Plonsky L	9	230	2010	5	9	0.454545

3.7   Highly Influential Research Journals

Top 10 most preferred journals that are publishing research on linguistics are presented in Table 4. There are five journals that produced over 100 publications. Notably, most of the research journals (04) belong to the Elsevier publisher. *Journal of Pragmatics* is on top of the list with 188 publications, remarkable 2977 citations against these publications, 15.83 citation impact and hence stand in Q2 ranked journal. The journal *Language Sciences* ranks second with 178 publications and 1363 citations with 7.65 citations impact. However, the rest of the journals though most preferred sources in the said list, have less than 150 publications. It is interesting to observe that majority of Linguistic researchers preferred good quality journals. There are four Q1 journals, three Q2, and only one journal of Quartile 3 and Q4, respectively.

Table 4: Source Impact

Rank	Title of Journal	TP	TC	IF	Q	Publisher	Country
1.	Journal of Pragmatics	188	2977	1.329	2	Elsevier	Netherlands
2.	Language Sciences	178	1363	0.853	3	Elsevier	UK
3.	Historiographia Linguistica (International Journal for the History of the Language Sciences)	125	239	0.375	4	John Benjamins Publishing Company	Netherlands
4.	Lingua	113	1097	0.963	2	Elsevier	Netherlands
5.	Applied Linguistics	109	3528	3.041	1	Oxford University Press	UK
6.	Linguistics	103	505	1.066	1	De Gruyter	Germany
7.	Journal of English For Academic Purposes	97	961	1.732	2	Elsevier	Netherlands
8.	Language	95	1995	1.899	1	Linguistic Society of America	Washington
9.	Eurasian Journal of Applied Linguistics	92	46	N.A	NA	Canakkale Onsekiz Mart University	Turkey
10.	Cognitive Linguistics	80	1416	1.630	1	De Gruyter	Germany

3.8   Authorship Pattern

The authorship pattern is shown in Figure 4. It is observed that joint collaborative research is happening globally; however, it is interesting that the single authorship pattern still dominates in the field of linguistic research. There are remarkable 3803 publications that are contributed by the single-author. Generally, the key authorship patterns are one, two, and three authors per publication. The maximum number of authorship pattern is thirty-authors.
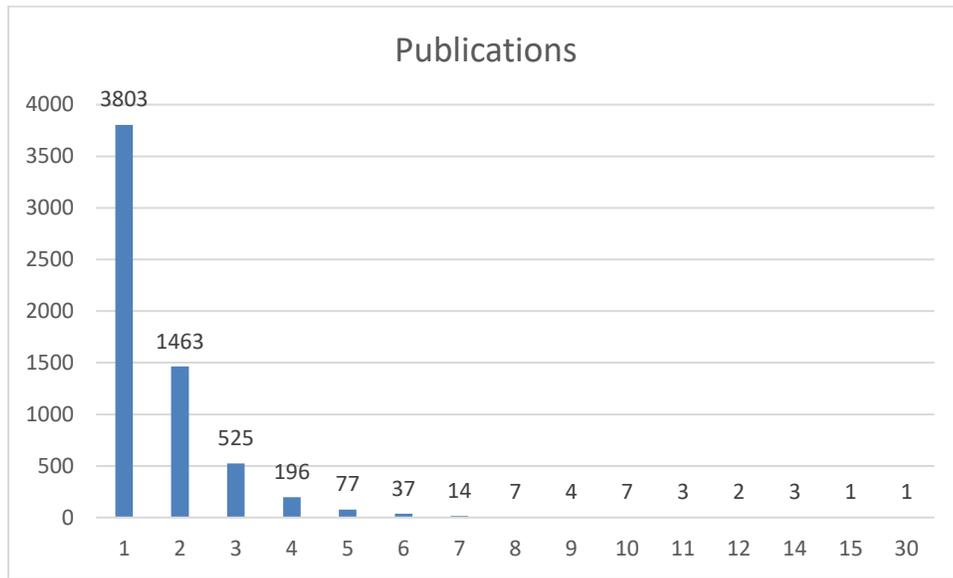

Figure 4: Authorship Pattern of Linguistic Researchers

3.9 Frequently Used Keywords

Frequently used authors' keywords in linguistic research are highlighted in Figure 5. The minimum number of 30 keywords occurrence is selected and hence only 53 keywords meet the threshold out of 11926 keywords. The distance and size of the bubble indicate number of the keyword occurrence and associational links. The top five keywords appeared more than 250 times. The keyword 'linguistics' is most frequently used as it appears 1702 times followed by the 'language' that appears 1094 times, 'English' appears in 518 publications, 'corpus' appears in 435 times, and 'discourse' that appears 279 times. VOS has generated six clusters of these 53 keywords. Cluster-

one (purple) has 16 keywords, including bilingualism, computational linguistics, conversational analysis, corpus linguistics, critical discourse analysis, discourse analysis, forensic linguistic, interactional linguistics, and sociolinguistics. Accordingly, other colors that are blue, orange, green, and pink are indicating associational links (Figure 5).
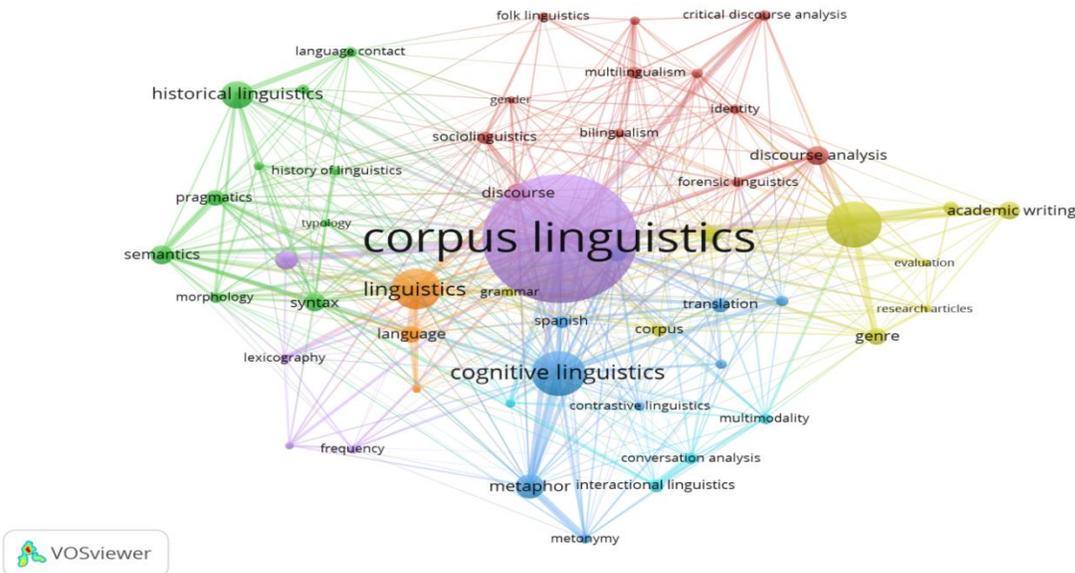
Figure 5: Frequently used keywords (1919-2020; minimum occurrence of words: 30)

Additionally, the list of most frequent keywords in the last 20 years to observe the latest trends in linguistic research is also generated. The result presented in figure 6 shows that "text linguistics, comparative linguistics, iconicity, sign language, is the main keywords that are repeated most frequently in linguistics literature from 2001 to 2019.
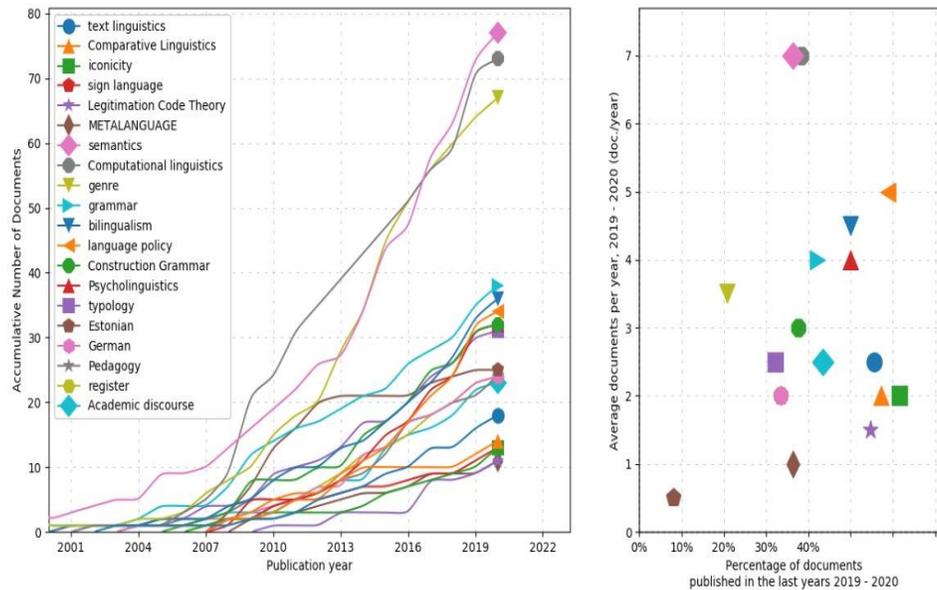
Figure 6: Author keyword analysis in two different chronological periods

3.10   Top Ten Highly Cited Articles

The bibliographic information of the top ten most cited articles is indicated in table 6. There are

five articles in this list that received over 100 citations. The publications years' range is between

2003 to 2016, and most of the articles in this list are published after 2010.The article entitled

"Language is a complex adaptive system" by ELLIS NC published in 2009 *in Language Learning*

is on the top of the list with 305 citations, followed by the article entitled "Language emergence"

by ELLIS NC in 2006 (Table 6). The top-cited article at the bottom of the list (Meta-analysis in

second language research: choices and challenges) is authored by 'Plonsky L' is published in 2010

in the *Annual Review of Applied Linguistics* have got 81 citations.

**Table 5:** Highly cited articles

Title	Author	Source title	Year	Citation
Language is a complex adaptive system	ELLIS NC	Language Learning	2009	305
Language emergence	ELLIS NC	Applied Linguistics	2006	213

Formulaic language in native and second language speakers	ELLIS NC	Tesol Quarterly	2008	161
Grammar is grammar and usage is usage	NEWMEYER FJ	Language	2003	150
A transdisciplinary framework for sla in a multilingual world	ELLIS NC	Modern Language Journal	2016	141
Converging evidence	GRIES ST	Cognitive Linguistics	2005	97
Study quality in SLA an assessment of designs, analyses, and reporting practices in quantitative l2 research	PLONSKY L	Studies in Second Language Acquisition	2013	87
Formulaic language and second language acquisition	ELLIS NC	Annual Review of Applied Linguistics	2012	82
Metonymy as a prototypical category	GEERAERTS D	Cognitive Linguistics	2006	81
Meta-analysis in second language research: choices and challenges	PLONSKY L	Annual Review of Applied Linguistics	2010	81

3.11 Three Factor Analysis

Keyword, Author and Source

The three-factor diagram of top 20 keywords, authors, and sources on linguistic literature has been generated. The size of the block shows the associational relationship with each factor. The top three authors (Crossley SA, Staples S, Parodig G) have a strong relationship by producing research on the top three sub-areas of linguistics (corpus linguistics, computational linguistics, linguistics) and prefer to publish in three sources (Modern Language Journal, International Journal of Corpus Linguistics, Journal of English for Academic Purpose).
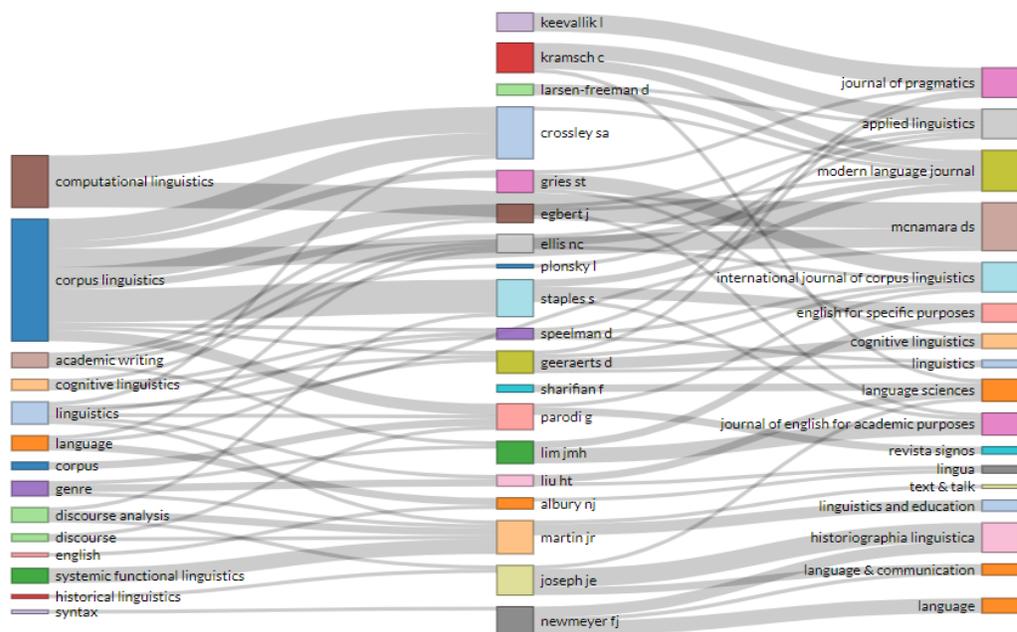
Figure 7: Three factor diagram of keyword (left), author (middle) and source (right)

Country, Author and Source

The three-factor diagram shows the relationship among top countries, authors, and their preferred platform for publications. Most of the top authors mainly belong to the US along with the authors from China, Belgium, and Australia. The top three authors (Ellis NC, Larsen-Freeman D, Gries ST, Crossly SA) have a strong relationship to publish in two sources (Modern Language Journal, Applied Linguistics). In addition, the top authors also prefer to publish their research in another three major journals (Histographia Linguistica, International Journal of Corpus Linguistics, Journal of Pragmatic).
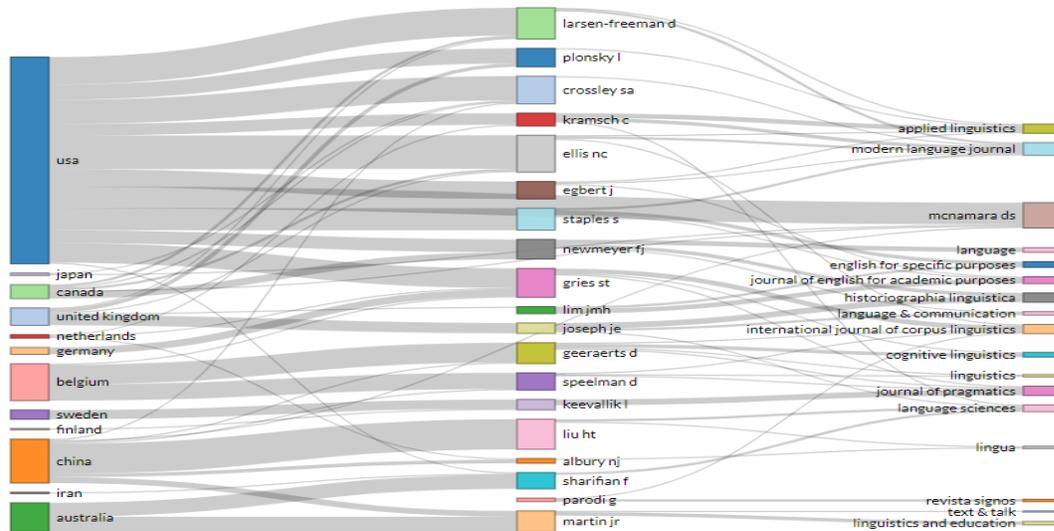
Figure 8: Three factor diagram of country (left), author (middle) and source (left)

4. Discussion

The results of the study provide useful insights regarding the research and productivity of linguistics during the last 100 years. It can be inferred from the results that an increase in linguistic publications is evident, especially a significant increase is observed after 1999 (see Figure 1). Region-wise research productivity has shown exponential growth in the US, the UK, and China, especially after 1995 (see Table 1). One possible reason may be the economic state of these countries where more funds are allocated for research for all disciplines as compared to other countries where funding for research projects is quite limited. The other reason may be that universities focus on research and its link with the promotion of university faculty. The findings of the study correlate with the Liu et al. (2015)'s research work, which in the context of a comparative study of four countries, reveals that economic growth plays a key role in the overall research progress of a country.

Overall, the increase in research publication obtained through results has been divided into two phases. The first phase (1919-1999) contributes only 9% of the total number of linguistics

publications. This implies that the publications in linguistics progressed at a low pace during the first phase. However, the second phase (2000-2020) in linguistics reveals exponential growth in research that is almost 91% of the total research counted in linguistics (see topic highly influential countries). This reflects a sudden surge in the publications and interest in linguistics research among major leading countries of the world. The results of the study are in partial agreement with the research conducted by Arık (2015), which shows an increase in social sciences and humanities publication, especially after 2005.

The findings of the study also reveal that the collaborative projects in linguistics are limited as compared to single authorship articles (see Figure 4). The results of the study are in line with previous studies (Ezema and Asogwa, 2014, Sharadha, 1991, Das et al., 2019) that conclude that single authorship dominates in linguistics and linguists for publication mostly rely on books and monographs as compared to research articles. The results are partially in line with Kumar and Kumar Kumar and Kumar (2011) research work, which concludes collaboration coefficient is higher in scientific research and for books and monographs, but in social sciences, the single-author pattern is preferred. This result implies that collaborative works in linguistics are limited, and the reason may be the lack of funding devoted to collaborative works.

The analysis of the frequently used key words reveals interesting information. It exhibits that emerging fields in linguistics have increased. These new fields include corpus linguistics, forensic linguistics, cognitive linguistics, translation, computational linguistics, etc. This implies that linguistics has not only grown up with the passage of time but also explored new avenues where it interacts with other disciplines such as in case of computational linguistics, it deals with computer and language, and in the case of forensic linguistics; it deals with the application of language in law (Amsler, 1982, Klein, 2015, Georgas and Cullars, 2005). The results also reveal

that the emergence of the new fields is linked especially to the second phase (see Figure 1). One possible reason is the core use of linguistics in disciplines like computer sciences, law, and psychology. One example of its application is natural language processing, which plays a key role in speech recognition, text analytics, topic modeling, etc. Similarly, in forensic linguistics, the application of language is evident, and linguistic analysis plays a supporting role in law. Finally, in the case of cognitive linguistics, which comes under psycholinguistics, the linguistic study reveals emotions that help psychologists reach at a conclusive stage (Casad, 2011).

The analysis of keyword, author, and source analysis reveals that computational linguistics and corpus linguistics have a strong connection with each other (see Figure 7). Authors whose interest area belongs to computational linguistics collaborate highly with authors in corpus linguistics. One possible reason is for most computational linguistics analysis, corpus-based data simplifies the task. (Souter and Atwell, 1993, Church and Mercer, 1993). Followed by computational linguistics and corpus linguistics in the area of academic writing, a higher degree of collaboration has been observed. One possible reason is that academic writing deals with the core aspects of writing linked to all disciplines (Hyland, 2004, Canagarajah, 2002).

Among the highly productive organizations, the University of Michigan is at the top of the list, followed by The French National Centre for Scientific Research (CNRS), France (see Table 2). The University of Australia is ranked after the above-mentioned institutions. An overall analysis of the results shows that most of the influential organizations belong to developed countries where research funding to institutions is the primary focus. One possible reason may be the focus on research productivity, quality of education, and a sufficient number of faculty in these institutions. The results of the study are in partial contradiction to the research conducted by Farooq and Al Shamrani (2020), which concludes that the dearth of faculty members, postgraduate

research programs, and quality of education is the main hurdle in research productivity in less developed countries.

The analysis of highly cited articles reveals that applied linguistics is the leading sub-fields in linguistic as most of the highly cited articles belong to the field of applied linguistics. Applied linguistics is the core field of linguistics, which centers on learning and teaching issues, but now its wider application encompasses the application of linguistic concepts in any field (see table 5). The results of the study are in line with the research work conducted by Lei and Liu (2019b), which reveals that research trends based on applied linguistics have increased, whereas research productivity in phonology, grammar, and generative linguistics has decreased. One possible reason may be its early emergence as compared to other subfields of linguistics and wider application linguistics addresses to the issues not only in language learning and teaching but also to relating acquisition of language and its status in different countries as a first or second language.

One possible reflection from the findings is the neglected areas in linguistics. The analysis of the second phase (200-2020) shows that the topics on meta-language, typology, iconicity, and comparative grammar have gained less attention (see Figure 6). Contrary to this, the interdisciplinary aspect of linguistics has taken more importance, as discussed above.

5.   Limitation and Future Research Directions

The study is limited only to articles indexed in the Web of Science database, and it is more likely that some more articles of linguistics indexed in other databases might have been missed out. Future researchers may focus on indexing of linguistic articles in other databases. On the other hand, no search query is 100% perfect, and false positive and false negative results are always a possibility. In certain cases, some authors or institutions might have more than one name or different name spellings. This might create an inaccuracy in the productivity of those authors or

institutions. Despite all these limitations, this study is the first comprehensive effort to analyze bibliometric indicators of linguistic literature.

Various future directions emerge from this study. One of the indicators from the research is the lack of collaborative projects. Research in linguistics should focus on collaborative projects. Similarly, research contributions are coming from the developed countries. The other countries in the world need to focus on this emerging discipline where language taxonomy is the core area. Findings of the study also indicate some of the neglected areas of linguistics research including comparative linguistics, typology, and language planning. It is the need of the hour to focus on these areas as research in these areas supports the comparison and promotion of languages.

6. Conclusion

The study aimed to investigate the publishing trends and patterns of linguistic literature in WOS. The major finding reveals that in the last two decades, research in linguistics has increased over time. Majorly the developed countries are in the leading position where linguistics is getting more focus of research. Moreover, in the case of collaboration between countries, the US has the highest number of collaborative projects as compared to other countries. The bibliometric analysis concludes that authorship patterns in linguistics mostly rely on single authorship as compared to collaborative projects. The trends show a single author pattern in linguistics as a dominant one. Therefore, collaboration in linguistic projects is on the basic level. The analysis of keyword author and source analysis sums up that computational linguistics and corpus linguistics have a higher degree of collaboration as compared to other areas in linguistics. Therefore, collaboration in linguistics can be summed up at two levels; intra-collaboration, inter-collaboration. The intra-collaboration level involves collaboration among sub-categories of linguistics and at the inter-collaboration level, it deals with interdisciplinary collaboration. The keyword analysis concludes that new fields in linguistics have emerged. These new subfields of linguistics are mostly the

outcome of the application of language in other fields such as computer sciences, laws, and psychology. Among these, computational linguistics, cognitive linguistics and forensic linguistics are more prominent. This is an indication of an increase in interdisciplinary research in linguistics that has taken a boost, especially in the last two decades. The analysis of the frequently used keywords brings another aspect to the fore, the neglected areas in linguistics. Metalanguage, comparative linguistics, typology and iconicity are the areas in linguistics where less attention has been paid. These areas have gained less attention, especially during the second phase (2000-2020).

Acknowledgements

References

AMERICA, L. S. O. 2020. *What is Linguistics?* [Online]. Available:
https://www.linguisticsociety.org/what-linguistics [Accessed].

AMSLER, R. A. Computational lexicology: a research program. Proceedings of the June 7-10,
1982, national computer conference, 1982. 657-663.

ARıK, E. 2015. A bibliometric analysis of linguistics in web of science.

CANAGARAJAH, A. S. 2002. *A geopolitics of academic writing*, University of Pittsburgh
Press.

CASAD, E. H. 2011. *Cognitive linguistics in the redwoods: The expansion of a new paradigm in
linguistics*, Walter de Gruyter.

CHING, G. 2012. 40 years of Language Studies: A Bibliometric Study of Research Articles in
Taiwan. *Senior Editors: Paul Robertson and Roger Nunn*, 53.

CHURCH, K. & MERCER, R. L. 1993. Introduction to the special issue on computational
linguistics using large corpora. *Computational linguistics,* 19**,** 1-24.

DAS, S., GAYAN, M. A. & VERMA, M. K. Authorship and Collaboration Pattern on Jaundice
Research Published by SAARC Countries during 1999-2018: A Scientometric Analysis.
First International Conference on "Science & Technology Metrics (STMet 2019)".
Pichappan Pit et. al.(Eds), Organized by Kind Mongkut's University of technology,
Bangkok, Thailand on, 2019. 02-04.

EZEMA, I. J. & ASOGWA, B. E. 2014. Citation analysis and authorship patterns of two
linguistics journals. *portal: Libraries and the Academy,* 14**,** 67-85.

FAROOQ, R. K. & AL SHAMRANI, F. J. 2020. Bibliometric analysis of Neurosciences
research productivity in Saudi Arabia from 2013-2018. *Neurosciences,* 25**,** 134-143.

GEORGAS, H. & CULLARS, J. 2005. A citation study of the characteristics of the linguistics literature. *College & Research Libraries,* 66**,** 496-516.

HYLAND, K. 2004. *Disciplinary discourses, Michigan classics ed.: Social interactions in academic writing*, University of Michigan Press.

JABER, S., AL-RAWAJFEH & AIMAN, E. 2017. Bibliometric analysis of the scopus-published articles on English Linguistics by Arab Scholars from 1970 to 2015. *Journal of Comparative Literature and Aesthetics,* 40**,** 112.

KLEIN, J. T. 2015. *Interdisciplining digital humanities: Boundary work in an emerging field*, University of Michigan Press.

KUMAR, S. & KUMAR, S. 2011. Citation analysis of journal of Oilseed Research.

LEI, L. & LIAO, S. 2017. Publications in linguistics journals from Mainland China, Hong Kong, Taiwan, and Macau (2003–2012): A bibliometric analysis. *Journal of Quantitative linguistics,* 24**,** 54-64.

LEI, L. & LIU, D. 2019a. The research trends and contributions of System's publications over the past four decades (1973–2017): A bibliometric analysis. *System,* 80**,** 1-13.

LEI, L. & LIU, D. 2019b. Research trends in applied linguistics from 2005 to 2016: A bibliometric analysis and its implications. *Applied Linguistics,* 40**,** 540-561.

LILJEGREN, H. 2018. Profiling Indo-Aryan in the Hindukush-Karakoram: A preliminary study of micro-typological patterns. *Journal of South Asian languages and linguistics,* 4**,** 107-156.

LIU, X., XU, Q. & LI, M. 2015. A comparative analysis of scientific publications in management journals by authors from Mainland China, Hong Kong, Taiwan, and Macau: 2003–2012. *Scientometrics,* 105**,** 135-143.

MOHSEN, M. A., FU, H.-Z. & HO, Y.-S. 2017. A Bibliometric Analysis of Linguistics

Publications in the Web of Science. *Journal of Scientometric Research,* 6**,** 109-118.

O'GRADY, W., DOBROVOLSKY, M. & KATAMBA, F. 1997. *Contemporary linguistics*, St.

Martin's.

REFERENCE, O. 2020. *Linguistics* [Online]. New York, NY 10016USA. Available:

https://www.oxfordreference.com/page/linguistics [Accessed].

SHARADHA, B. 1991. Multiple-authorship pattern in linguistics. *SRELS Journal of Information

Management,* 28**,** 96-99.

SOUTER, C. & ATWELL, E. 1993. *Corpus-based computational linguistics*, Rodopi.

WEINREICH, M. 2010. Language Shift in Northern Pakistan: The Case of Domaakí and Pashto.

Iran and the Caucasus, 14**,** 43-56.