LingBaW *LingBaW. Linguistics Beyond and Within 11 (2025), 119–140*

Back to Normality? A Bibliometric Analysis of L2 Vocabulary Research 1989–93

Paul Meara
Swansea University and University of Oxford, UK

Abstract

This paper is the latest in a series of studies published in *LingBaW* that contribute towards a historical review of research into L2 vocabulary acquisition. This instalment uses an author co-citation analysis (Small: 1973) to examine the research on L2 vocabulary acquisition published in 1993. Two analyses are presented. The first analysis provides a context for the 1993 data. It looks at work that was being cited in a five-year window covering 1989–93. The second analysis is a more detailed account of the 1993 research on its own terms. These analyses suggest that a surge in psycholinguistic research identified in 1992 may be a transient feature of the field. 1993 exhibits a return the patterns noted in our earlier studies, where the most significant influences come from an applied linguistics perspective.

Keywords: L2 vocabulary acquisition; vocabulary research; bibliometric; author co-citation

1. Introduction

This paper is the thirteenth instalment in a series of studies in which I have been attempting to map out the development of L2 vocabulary research over the last 50 years (cf. Meara 2024). These reports are based on the research outputs listed in the Vocabulary Acquisition Research Archive (VARGA) database (Meara n.d.), a comprehensive collection of research papers that has become the main reference source for research into L2 vocabulary acquisition. Previous reports covered the period 1982–92. The present report takes this historical overview another step forward by analysing the new research which appeared in 1993.

In my last report (Meara 2024), I noted that the 1992 outputs amounted to a marked change of direction from my earlier reports. The picture that had been developing since 1982 showed a steady growth in the number of research outputs, and the emergence of identifiable clusters in the bibliometric maps that signal developing trends in the L2 vocabulary research. In 1992, however, these trends were seriously disrupted: the 1992 bibliometric map was no longer dominated by influencers from the tradition of Applied Linguistics. Rather, the most-

LingBaW. Linguistics Beyond and Within, e-ISSN: 2450-5188
DOI: 10.31743/lingbaw.19184

cited influencers in 1992 all came from a more psycholinguistic tradition, asking different questions and relying on highly technical experimental research methods which had previously played only a limited and minor role. What we might call "main stream" L2 Vocabulary Research was reduced in 1992 to a small cluster exhibiting few co-citation links with the much larger psycholinguistics cluster.

At the time, I questioned whether this dramatic change of course was a permanent feature of the research, or just a temporary blip. A preliminary assessment of the 1993 data suggests that "normality" has been restored: the burst of activity in L2 psycholinguistics has retrenched and themes identified in our earlier studies have re-emerged. Indeed, as we will see, most of the new influencers who appeared in the 1992 mapping have disappeared in 1993, being replaced by more familiar names.

This report begins with an overview of the research published in the five-year window 1989–93, and continues with a more detailed exploratory account of the 1993 publications. As usual, the analyses that follow use the Author Co-citation method developed by Small (1973). Small's methodology is described in detail in Appendix 1 for the benefit of readers who are not yet familiar with the approach used in these reports.

2. Part 1: The 1989–93 data set

The basic statistics of the 1989–93 data set are summarised in Table 1, alongside the 1988–92 data for comparison. The table shows that the 1989–93 corpus is considerably larger (17%) than the 1988–92 corpus. We can also identify a significant increase in the number of authors who contribute to the data set. As usual, most of the authors identified make just one contribution to the corpus: the proportion of authors in this category remains steady at about 81% in 1989–93. The number of prolific authors – authors who contribute six or more outputs to the data set - has slightly increased in 1989–93.

Table 1: *The basic statistics of the 1988–92 and the 1989–1993 research outputs*

	1988–92	**1989–93**
Total outputs	628	734
Unique authors	633	731
Prolific authors (6+ contributions)	12	15
Authors making a single contribution	512	573

Table 2: *The prolific authors in the 1988–92 and 1989–1993 research outputs. (Prolific here is defined as a contribution to at least six outputs)*

	1988–92	**1989–93**
10+	Meara (21) Laufer (18) Vermeer (10)	Laufer (20) Meara (20) Vermeer (11) Nation (10)
9		Arnaud
8	Broeder	
7	Carter Johns McCarthy Nation Zimmerman	Appel Broeder
6	Appel Arnaud Palmberg	Abe Bogaards Decoo Hartmann Hulstijn Schouten-van Parreren Weltens Zimmerman

Table 2 lists the most prolific authors in the 1989–93 data. This prolific author list is slightly larger than the equivalent list for 1988–92 – a reflection of the general increase in the number of outputs appearing in each successive year. Four authors have dropped out of this list (Carter, McCarthy, Johns and Palmberg), and seven new authors have joined the list (Abe, Bogaards, Decoo, Hartmann, Hulstijn, Schouten-van Parreren and Weltens. Abe was based in Japan, and hints at a growing Japanese interest in Vocabulary Acquisition research. His work is mostly concerned with the development of communicative vocabularies. Bogaards and Hartmann represent the dictionary use strand of research. Decoo's work is mainly concerned with CALL and vocabulary. Weltens' papers all deal with vocabulary loss – a new theme to appear in this series of maps. Schouten-van Parreren and Hulstijn are mainly concerned with how L2 readers handle unknown words.

The analysis that follows uses the author co-citation method developed by Small (1973) (see Appendix 1). By convention, not all outputs are included in author co-citation analyses. Book chapters and papers published in journals are included, but other types of output (monographs, theses, computer programs, and so on) are not. The rationale for these exclusions is that book chapters and journal papers tend to have a consistent approach to citation of other people's work while the other types of output often take a different approach to citing their sources. These different practices distort the statistical trends in the data. Theses, for example, usually reference enormous numbers of sources, whereas journal papers are typically more sparing in their approach. The next step in our analysis therefore involves pruning the corpus to generate a smaller data set that consists of journal articles and book chapters (the *eligible data set)*. The results of this pruning process are reported in Table 3. The table shows a modest increase in the number of eligible outputs in the 1989–93 data set, and a similarly small increase in the number of authors contributing to the data set. The number of unique authors who are cited in this dataset is 6055 – an increase of 16% over the 1988–92 figure of 5210, a figure that we previously noted was already very large. As usual, the authors cited only once make up the largest part of the data (3920 cases). The complete distribution is shown in Table 4.

Table 3: *The main characteristics of the 1988–1992 and the 1989–1993 eligible data sets*

	1988–92	**1989–93**
Number of outputs in the eligible data set	464	542
Number of authors contributing to the eligible data set	421	743
Number of influencers cited in the eligible data set	5210	6055

Table 4: *The number of cases cited N times in the 1989–93 data set*

FREQ	**75+**	**74**	**73**	**72**	**71**	**70**	**69**	**68**	**67**	**66**	**65**	**64**	**63**	**62**	**61**
Cases	2			1			1								
FREQ	**60**	**59**	**58**	**57**	**56**	**55**	**54**	**53**	**52**	**51**	**50**	**49**	**48**	**47**	**46**
Cases					2					1		1		1	
FREQ	**45**	**44**	**43**	**42**	**41**	**40**	**39**	**38**	**37**	**36**	**35**	**34**	**33**	**32**	**31**
Cases	2	1	1			1	2	1	2	4	3	2	3	5	
FREQ	**30**	**29**	**28**	**27**	**26**	**25**	**24**	**23**	**22**	**21**	**20**	**19**	**18**	**17**	**16**
Cases	4	6	3	4	2	1	5	8	9	11	4	12	8	13	12
FREQ	**15**	**14**	**13**	**12**	**11**	**10**	**9**	**8**	**7**	**6**	**5**	**4**	**3**	**2**	**1**
Cases	14	21	31	22	28	31	39	69	64	103	137	210	377	851	3920

Table 5 lists the most frequently cited sources in this data set. This table shows that the number of authors cited at least 40 times has increased in the 1989–93 data set. Nation and Laufer both show very large gains: Nation's citations have increased by 26%, while Laufer's citations have increased by 50%. Sinclair has dropped out of the list, but in his place we have seven new entrants: RC Anderson, Schouten-van Parreren, Gairns, Redman, Nagy, McCarthy and AD Cohen.

Table 5: *The most frequently cited authors in the 1988–92 and the 1989–93 data sets*

1989–92	1989–93
Nation (76)	Nation (102)
Meara (74)	Meara (83)
Carter (54)	Laufer (72)
Krashen (49)	Carter (69)
Laufer(48)	Krashen, Richards (56)
Richards (46)	RC Anderson (51)
Sinclair (44)	Schouten-van Parreren (49)
Lockhart (43)	Lockhart (47)
	Gairns, Redman (45)
	Nagy (44)
	McCarthy (43)
	AD Cohen (40)

Table 6: *The principal statistics of the 1988–92 and the 1989–93 eligible data sets*

	1988–92	1989–93
Inclusion Threshold	16	19
Sources included	103	105
New Influencers		20
Lost Influencers	18	

The analysis that follows is based on the co-citations among the most frequently cited authors in the 1989–93 data set. It is normal practice in author co-citation studies to work with the 100 or so most frequently cited authors. The data in Table 4 suggest that we can get close to this conventional figure if we adopt an inclusion threshold of 19 citations in the data set. This threshold identifies a set of 105 highly cited authors to work with. For simplicity, we will refer to this set as a list of influencers. In our analysis of the 1988–92 data set, 103 influencers met an inclusion threshold of 16 citations. The 1989–93 data set is therefore very comparable to the 1988–92 data set in terms of size, but slightly more demanding in its threshold (138 influencers in the 1989–93 data set meet the threshold we used in our 1988–92 analysis with 16 citations or more). These characteristics are summarised in Table 6.

Again, the rate of churn between the two data sets is relatively small. 18 highly cited authors that appear in the 1988–92 dataset no longer meet our inclusion threshold (Bejoint, D Clarke, Coltheart, R Ellis, Greenbaum, Hartmann, Hornby, Johns, D Meyer, Olshtain, Putseys, Renouf, Ruddy, Rudzka, Sharwood-Smith, Wallace, Widdowson and Zimmerman), while 20 new entrants appear as Influencers in the 1989–93 data set (JR Anderson, Baddeley, Carrell, Coady, Cummins, Delaney, de Groot, Freebody, Johansson, Kroll, Lorge, McLaughlin,

Morgan, Obler, Rinvolucri, Schvaneveldt, NS Segalowitz, Seidenberg, Stanovich and Sternberg). These cases will be discussed in Section 3.

The co-citations among the 105 influencers in the 1989–93 data set were mapped using the Gephi software package (Bastian, Heymann & Jacomy 2009). Co-citations which appear only once in the data set are ignored: Gephi identifies 2728 co-citation links that appear in the data set at least two times. In the analysis that follows, I asked Gephi to generate a spanning tree map, based on the strongest co-citations within the data set. The methodology for building spanning trees is explained in more detail in Appendix 3. The spanning tree approach is less informative than the complete mappings that I presented in earlier reports but it has a number of advantages – the obvious advantage being that the spanning trees are much easier to read and interpret than the complete maps. A more subtle advantage is that the maps in my earlier reports were based on arbitrary inclusion thresholds, and these made it difficult to compare a map based on one set of data with another map derived from a different set of data. Working with spanning trees avoids these arbitrary criteria: each spanning tree maps the most cited influencers in a particular data set. Spanning trees for successive time periods are more directly comparable with each other, and changes over time are easier to identify in this format.

Figure 1 shows the basic mapping of the 1989–93 data set. This map shows the 105 influencers who are cited at least 19 times in the data set. Each influencer appears as a node in the map; the size of a node indicates how many other nodes it is connected to.

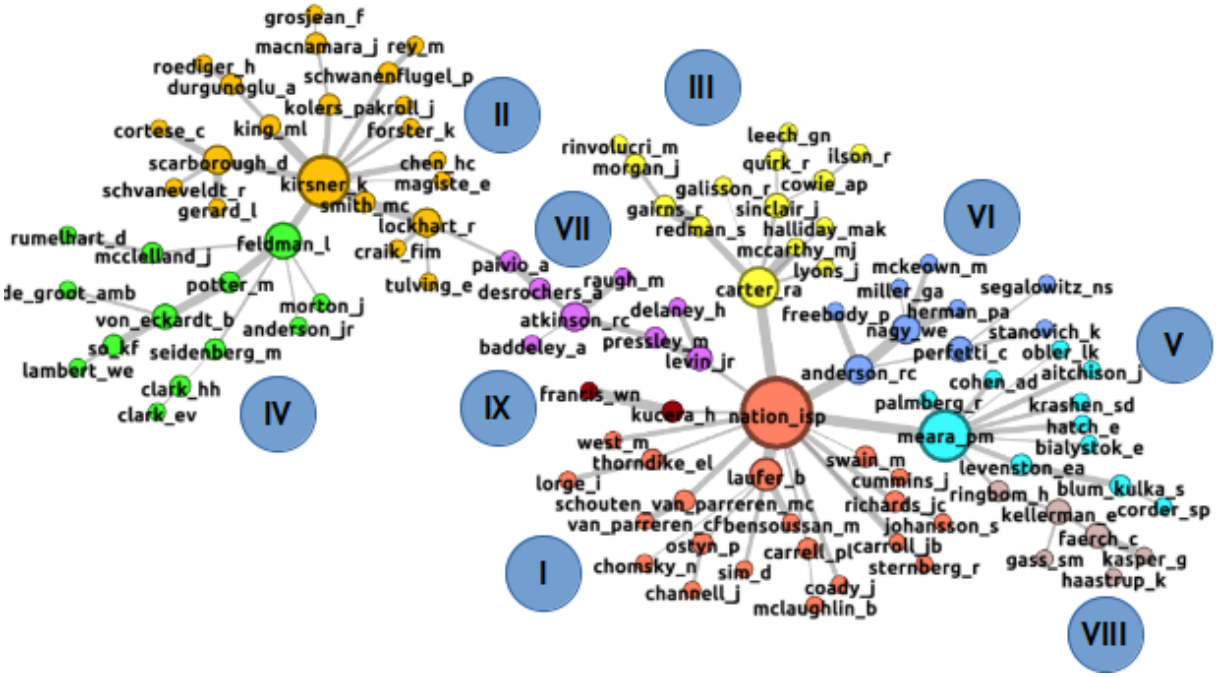

Figure 1: *A spanning tree map of the 1989–93 data set. The map contains 105 nodes with at least 19 citations in the data set. Colours indicate the nine thematic clusters identified by Gephi. Nodes are sized according to how many connections they have with other nodes.*

Gephi identifies nine clusters in this map.

The largest cluster, **Cluster I** (21 members), is dominated by Nation, who replaces Meara as the central node in this map. We can identify a word frequency theme in this cluster (West

and Thorndike & Lorge) and also a sub-cluster centred on Laufer. This cluster is larger than the main cluster that we identified in the 1988–92 mapping, despite some of the subclusters developing into clusters in their own right.

Cluster II, centred on Kirsner (15 members) is essentially the same group of influencers that appeared as Cluster III and cluster VII in the 1988–92 map. A significant addition to this cluster is Judith Kroll.

Cluster III (14 members), focused on Carter and Sinclair, is a group of linguists whose main interests lie in descriptions of English and corpus linguistics. This group of influencers also appeared in the 1988–92 mapping, but here it also includes Gairns & Redman and Morgan & Rinvolucri – influencers with a strong and direct interest in vocabulary pedagogy.

Cluster IV (13 members) focussed on Feldman, is essentially the same as cluster VI in the 1988–92 map, with the important additions H Clark and EV Clark (L1 vocabulary acquisition) and of AMB de Groot.

Cluster V (11 members focussed on Meara) makes up the remnants of Cluster II in the 1988–92 map. This cluster is much less central than it was in last year's mapping, largely due to the loss of the Israeli sub-cluster, which has now become part of Cluster I.

Cluster VI (9 members) is a new cluster focussed on Nagy and RC Anderson. This cluster mainly consists of L1 reading specialists.

Cluster VII (8 members) is a group of psychologists whose main focus is the use of imagery and mnemonics for learning vocabulary. A significant addition to this group is Baddeley, whose work on memory is an important emerging theme in this map.

Cluster VIII (6 members) is a group of predominantly Scandinavian researchers, mainly interested in lexical transfer.

Gephi also identifies **Cluster IX**, a cluster consisting of two members – Kucera and Francis. This frequency count is probably best seen as part of the frequency count subcluster that we identified in Cluster I.

Table 7 summarises the main features of this map, and provides comparison figures for the equivalent map covering the 5-year window 1988–92.

Table 7: *The main thematic clusters in the 1988–92 data set*

cluster	1988–92	1989–93
I	Vocabulary learning theory (18)	Vocabulary teaching and reading (21)
II	Vocabulary teaching and reading (17)	Bilingual Word recognition (21)
III	Bilingual word recognition (15)	Corpora, Discourse and Dictionaries (14)
IV	Corpora and Discourse (12)	Performance of bilinguals (13)
V	Lexical error and transfer (9)	L2 vocabulary learning theory (11)
VI	Psycholinguistics (9)	L1 reading (8)
VII	Performance of bilinguals (7)	Imagery, Mnemonics and Memory (8)
VIII	Imagery and Mnemonics (6)	Lexical error and transfer (6)
IX	Dictionaries and their use (5)	Kucera & Francis (2)
X	Applications of Semantics (4)	

The two maps are actually very similar: the thematic clusters identified in the 1988–92 map are easily recognizable in the 1989–93 map, but some subtle shifts in the structure of the field

can also be identified. The main change is that the biggest cluster is now dominated by Nation. The largest cluster in the 1988–92 map has fractured into two smaller clusters (Cluster V and Cluster VIII). Some of the smaller clusters in the 1998–92 map have coalesced into larger clusters (Cluster IV and Cluster III in the 1989–93 map). A set of influencers dealing with L1 reading has emerged as a cluster in its own right (Cluster VI). A number of other minor changes can also be identified: there is some growth in the imagery and mnemonics cluster (Cluster VII); the L1 acquisition subcluster (Clark & Clark) has relocated to the psycholinguistics cluster (Cluster VI); the lexical transfer cluster seems to have shrunk; and Ostyn & Channell, the only survivors of cluster X in 1988–92, have been absorbed into the main L2 vocabulary pedagogy cluster (Cluster I). The bilingual word recognition cluster that we identified as an anomaly in the 1988–92 map seems to have grown in importance, and now makes up Cluster II in the 1989–93 map.

The 1989–93 map seems to fall into four main sectors. Clusters II and IV are methodologically reliant on experimental studies of bilingual speakers. Cluster III uses applied linguistics methodology, particularly corpus analysis. Cluster VI stands out as a set of L1 reading specialists whose work has strongly influenced L2 vocabulary research, and is particularly closely associated with Cluster I. The remaining clusters are linked by a methodological concern with L2 learners.

Table 8: *The strongest co-citation links in the 1988–92 and the 1989–93 data sets. Persistent links that appear in both lists are shown in* ***bold****.*

1988–92	1989–93
Kirsner ~ Smith 32	**Gairns ~ Redman 45**
Carter ~ McCarthy 32	**Carter ~ McCarthy 40**
Gairns ~ Redman 32	**Laufer ~ Nation 37**
Kirsner ~ Lockhart 30	Anderson ~ Nagy 36
King ~ Smith 29	**Kirsner ~ Smith 34**
Meara ~ Nation 29	Kucera ~ Francis 33
Laufer ~ Nation 27	**Meara ~ Nation 33**
Faerch ~ Kasper 26	Nagy ~ Herman 32
	Lockhart ~ Perfetti 32
	Carter ~ Nation 32

Table 8 lists the strongest co-citation links in this data set. The pattern here broadly reinforces the features that we noted in Figure 1. Most of the strong co-citations that we identified in the 1988–92 data set persist into 1989–93, with a particularly strong showing from Gairns & Redman. Of the new entries in the 1989–93 list, Anderson~Nagy and Herman~Nagy signal the increasing influence of L1 reading studies for L2 reading, while the Carter~Nation link underlines the continuing influence of corpus studies on L2 vocabulary research. A number of strong links in the 1988–92 data fail to make it into the 1989–93 list. Two of these (Kirsner~Lockhart and King~Smith) contributed to the importance of the Bilingual Word Recognition cluster in the 1988–92 data set. Their loss hints that the prominence of this cluster may be temporary. More important is the loss of the Faerch~Kasper co-citation, which was a significant feature of the Scandinavian research cluster.

3. Part 2: The 1993 data in more detail

We now turn to a more detailed analysis of the research published in 1993.

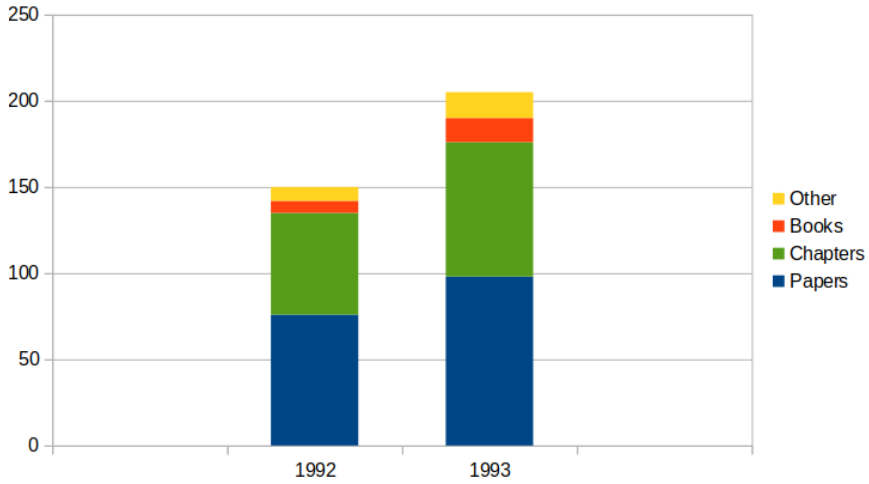

Figure 2: *The number of outputs published in 1992 and 1993 categorised by output type*

Figure 2 shows the publications that appeared in 1993 by output type, alongside the equivalent data for 1992. The obvious point to make here is that 1993 shows a large increase in the number of outputs, with an especially marked increase in the number of papers that appear as book chapters. The number of book length treatments appearing in 1993 is about twice the number that appeared in 1992. The fourteen 1993 items are listed in Table 9.

Table 9: *The fourteen book length items that appeared in 1993*

Arnaud, PJL and P Thoiron. *Aspects du vocabulaire.*[Aspects of Vocabulary.] Lyon: Presses Universitaires de Lyon. 1993.
Börner, W and K Vogel (eds.). *Wortschatz und Fremdsprachenerwerb.*Bochum: ASK. 1993.
Chapelle, J and M Claes (eds.). *Proceedings of the first international conference on memory and memorization in acquiring and learning languages.* Louvain la Neuve. 1993.
Huckin, T, M Haynes and J Coady (eds.). *Second Language Reading and Vocabulary.* Norwood, NJ.: Ablex. 1993.
Schreuder R and B Weltens (eds.). *The Bilingual Lexicon.* Amsterdam:John Benjamins. 1993.
Cribbin, L. 150 × false friends: typische Wortschatzfehler deutsch–englisch. München. 1993.
Kühnel, H. Typische Fehler Italienisch (3000 "falsche Freunde" italienisch und deutsch). Berlin 1993.
Leiste, D, C Döll and AM Tereso Domingos. *Typische Fehler Portugiesisch und Deutsch.*Berlin. 1993.
Daams-Moussault, A. *Nouveau vocabulaire de base PLUS.*Kerkrade: Uitgeverij NIB. 1993.
Harrison, WF and D Winters Welker. *Spanish Vocabulary Book: a new approach to vocabulary building.* Austin: University of Texas Press. 1993.
Lewis, M. *The Lexical Approach.*Hove: Language Teaching Publications. 1993.
Löschmann, M. *Effiziente Wortschatzarbeit.* [Efficient vocabulary training.] Frankfurt/Main: Lang. 1993.
Picoche, J. *Didactique du vocabulaire français.*Paris: Nathan. 1993.
Panzer, B. (ed.) Aufbau, Entwicklung und Struktur des Wortschatzes in den europäischen Sprachen. Motive, Tendenzen, Strömungen und ihre Folgen. Frankfurt/M.: Lang, 1993.

The first five items are thematic collections of chapters that report work on vocabulary acquisition. Most of this work was first presented at conferences that took place in 1991 or

1992. The next three items are collections of False Friends in specific language pairs. **Daams-Moussault** is a pedagogical word list for L1_Dutch learners of French. **Harrison & Winters Welker** is mainly concerned with the use of mnemonics for learning a 2000-word Spanish vocabulary. **Lewis** is an important pedagogical statement about the importance of vocabulary in language learning that will become Increasingly significant in future years. I was unable to obtain a copy of **Löschmann** due to ongoing issues at the British Library. This book appears to look at the psychological and linguistic theory that underpins currently used vocabulary exercises, with a particular emphasis on learning strategies, teaching strategies, semanticisation techniques and vocabulary tests. The final chapter proposes some new pedagogical approaches exploiting the possibilities afforded by new media.

Table 10: *Items published in 1993 that fall into the Other category*

Theses

Al-Hazemi, HA. *Low-level EFL vocabulary tests for Arabic speakers.* PhD Thesis. University of Wales. 1993.

Grendel, M. Verlies en herstel van lexikale kennis. [Loss and recovery of lexical knowledge.] Doctoral Dissertation, Katholieke Universiteit Nijmegen. 1993.

Hasselgren, A. Right words, wrong words and different words. Masters Thesis. University of Bergen. 1993.

Luo J. *A study of the effects of marginal glosses on the reading comprehension of intermediate college students of French.* Doctoral dissertation. University of Pennsylvania, Philadelphia, PA. 1993.

Mol, M. *Deducing words from context: a video-taped training project in reading strategies for the English central examination.* Doctoraalscriptie: Vrije Universiteit Amsterdam. 1993.

Newton, J. The relationship between pedagogic tasks, interaction and language learning. PhD thesis: Victoria University of Wellington. 1993.

Todd, HF. The effectiveness of multimedia technology in the acquisition of Spanish vocabulary. PhD Thesis. University of Pennsylvania. 1993.

Zhang, X. English collocations and their effect on the writing of native and non-native college freshmen. PhD Thesis. Indiana University of Pennsylvania. 1993.

Other formats

Arnaud, PJL. Mille lapsus. Unpublished manuscript CRTT, Université Lumière-Lyon 2. 1993.

Daams-Moussault,A and FMM Blaauw-Holtzappel. *DocentenBrochure: Nouveau vocabulaire de base, Nouveau vocabulaire de base PLUS, Nouveau vocabulaire francais, Livrets de controle.* Kerkrade: Uitgeverij NIB. 1993.

Qian Dawei. Teaching and learning vocabulary in a second language. An annotated bibliography. Ontario: OISE. 1993.

Toya, M. Form of explanation in modification of listening input in L2 vocabulary learning Occasional Paper 23. Honolulu: Department of ESL, University of Hawaii at Manoa. 1993.

Tréville, M-C. Role des congénères interlinguaux dans le developpement du vocabulaire réceptif. Quebec: Centre International de Recherche en Aménagement Linguistique. 1993.

Watanabe, Y. Effects of increased processing on incidental learning of foreign language vocabulary. Occasional Paper 24. Department of ESL, University of Hawai'i at Manoa. Honolulu. 1993.

Yang, L and T Givon. Tracking the acquisition of L2 vocabulary: the Keki language experiment. Eugene Or.: Institute of Cognitive and Decision Sciences. University of Oregon. Technical Report 93–11. 1993.

Picoche is a companion volume to her 1992 book (Picoche 1992) but the 1993 text deals more explicitly with L2 pedagogy. **Panzer** is principally concerned with loan words from the Slavic languages into German, and touches only rarely on the implications for L2 learners of German.

The **Other** category in Figure 2 is mainly composed of theses and unpublished reports. The eight theses recorded for 1993 are listed in Table 10. All these theses are cited in later publications, but readers should note that the VARGA database does not routinely monitor publications of this type, and the actual number of theses that appeared in 1993 is probably larger than this list of eight suggests. A small number of other output types also appeared in 1993. **Arnaud** is a collection of 1000 (mostly lexical) errors made by L1 French learners of English. **Daams-Moussault** is a teachers' handbook that accompanies the word list described in Table 9. **Qian** is an annotated bibliography. **Toya, Treville, Watanabe** and **Yang & Givon** are unpublished pieces that I was not able to obtain a copy of.

A total of 249 individual authors can be identified in the 1993 data set. Table 11 lists the number of authors according to the number of outputs that they contribute to. The most important feature in this table is that the number of contributing authors has increased significantly compared with 1992. In 1992, we identified 185 contributing authors; in 1993, this figure has risen to 249, a increase of just over 20%. As usual, most authors (205) contribute to only a single output. However, the number of authors contributing to two outputs has almost doubled, and there is a very significant increase in the number of authors contributing to three outputs (only one in 1992, but nine in 1993). This year, the most prolific author is Paul Nation (five outputs). Haynes (a new entry in the prolific author lists) contributes to four outputs. The nine authors contributing to three outputs in 1992 are: Arnaud, Coady, Huckin, Laufer, Meara, Schouten-van Parreren, Schreuder, Vermeer and Weltens. The prominence of Dutch researchers is striking here.

Table 11: *The number of authors contributing to N outputs in 1993*

N outputs	8	7	6	5	4	3	2	1
Cases in 1993				1	1	9	33	205
Cases in 1992				1	2	1	18	163
Lotka's estimate			6	8	13	23	51	

Table 12 lists all the prolific authors who published in 1993. (Here "prolific" means more than two contributions to the data set.) Only Arnaud, Laufer, Meara and Vermeer have a presence in both lists: the majority of the prolific authors identified in 1992 no longer qualify for inclusion in the 1993 list, and the majority of the 1993 prolific authors are new. Three other features in the 1993 list are worth noting. Only five of the prolific authors from 1992 have any kind of presence in the 1993 data set: Schouten-van Parreren is a familiar figure from earlier years; Hulstijn contributed to two outputs in the 1993 data set; de Groot, Grainger and Thomas & Wang contributed to a single output in 1993. The two most prolific authors in 1993 (Nation and Haynes) have not appeared in our previous prolific author lists. The appearance of edited collections that deal specifically with L2 vocabulary topics is a new and important development in the field; of the new appearances in the list, Huckin, Haynes & Coady and Schreuder & Weltens are editors of books of this type (cf. Table 9). Lotka's estimate is discussed in Appendix 3.

Table 12: *The prolific authors in 1992 and 1993. (Here, "prolific" means more than one contribution to the data.) The majority of the authors who counted as prolific in 1992 are no longer listed as such in 1993. Authors appearing in both lists are shown in* ***bold****.*

1992	1993
Laufer, Meara, Arnaud, Bejoint, de Groot, Grainger, Harrington, Hartmann, Hulstijn, Leffa, Löschmann, Verneer, Doctor & Klein, McLaughlin & Heredia, Thomas & Wang, Pearson, Umbel, Oller & Fernandez	Nation, **Arnaud, Laufer, Meara, Vermeer,** Schouten-van Parreren, Schreuder & Weltens, Huckin, Haynes & Coady,

3.1. The data sources

The VARGA data base (Meara n.d.) identifies 176 items published in 1993 that were eligible for inclusion in the analysis that follows. Fifteen of these outputs were not traceable, mostly because of a data breach at the British Library that led to temporary suspension of the British Library On Demand service. These outputs are listed in Table13. Most of these items were chapters in books that were not obtainable in other UK libraries. The remaining 162 outputs (88 journal articles and 74 book chapters) make up the data set that is analysed in the report that follows. For space reasons, I have not listed all these items in the report. However, interested readers can identify these included items by accessing the VARGA database at https://www.lognostics.co.uk/varga and entering the search terms 1993 **JA}** and 1993 **CH}.** Code JA} identifies journal articles, and CH} identifies book chapters.

Table 13: *Items published in 1993 that were not traceable*

Beheydt, L. Vocabulaireverwerving. Een alternatieve visie. In: R Trampus-Snel (ed.), *Nederlandse Taal-, Vertaal- en Letterkunde*. Trieste. 1993.

Hausmann, FJ. Ist der deutsche Wortschatz lernbar? Oder: Wortschatz ist Chaos. *Info DaF* 20,5(1993), 471–485.

Kątny, A. "Falsche Freunde" in den deutsch-polnischen Beziehungen. In: A Kątny (ed.) *Beiträge zur Sprachwissenschaft: Sozio- und Psycholinguistik. Probleme des Deutschen als Mutter-, Fremd- und Zweitsprache.* Rzeszów: 1993. 55–67.

Lipczuk, R. Faux Amis, Tautonyme, Internationalismen. *Studia i materiały: Germanistyka* 10. Zielona Góra 1993, 29–38.

Mazière, F. Le mot, unité didactique : une entrée dans la langue par le mot. *Repères* 8(1993), 29–39.

Meißner, FJ. Zukunftsmusik? - Überlegungen zu einem elektronischen Lernerwörterbuch mit französischen, spanischen und italienischen Beispielen. *Fremdsprachenunterricht* I(1993), 43–46 and II(1993) 104–106.

Nienhuis, LJA. Oefenvormen voor vocabulaireverwerving. In: A Toussaint-Dekker (ed.) *Verder met Frans.* 's-Hertogenbosch: Katholiek Pedagogisch Centrum (KPC). 1993.

Schatte, C. Repräsentanten von Internationalismen im Polnischen unter dem Aspekt ihrer Schreibung und Lautung. In: LM Eichinger and J Raith (eds.) *Sprachkontakte: Konstanten und Variablen.* Essen. 1993. 173–180.

Schiffler, L. Faux amis. *Fremdsprachenunterricht* 37,8(1993), 486–487.

Sheshsha, JA. Lexical error analysis in learning English as a Foreign Language. *Social Science Research Series. Umm Al-Qura Univesity.* Makkah, Saudi Arabia, 24(1993), 5–30.

Tamamura, F. Nihongo ni okeru kanji: sono tokushitsu to kyōiku. *Nihongo kyōiku* 80(1993), 1–14. **Verhallen, M** Kennis van woordbetekenissen bij tweetalige kinderen. *Psychologie en Maatschappij* 17,(1993), 129–146.

Vermeer, A. Woordenschatverwerving in het onderwijs. [Vocabulary development in teaching.] *VONK, Tijdschrift van de vereniging voor het onderwijs in het Nederlands* 22,5(1993) 15–22.

Volmert, J. Internationalismen – unter Aspekten des Fremdsprachen- und Zweitsprachenerwerbs. In: A Katny (ed.) *Beiträge zur Sprachwissenschaft: Sozio- und Psycholinguistik. Probleme des Deutschen als Mutter-, Fremd- und Zweitsprache.* Ed. Andrzej Kątny. Rzeszów 1993. 67–82.

3.2. The analysis

The first step in our analysis is to identify the authors who are cited most frequently in this data set, the Significant Influencers. These data are summarised in Table 14. The table shows that a total of 2860 influencers are cited with 2065 influencers being cited only once in the data set. The equivalent figures for 1992 are 2033 and 1511: the figures indicate that the number of influencers cited in the 1993 data set has increased by about 40%, with the proportion of authors cited only once remaining stable at 74%. The most frequently cited influencers in the 1993 data set are Nation (cited in 41 outputs), Laufer (cited in 29 outputs), Carter (cited in 25 outputs), RC Anderson (cited in 24 outputs), Meara (cited in 23 outputs), Nagy (cited in 21 outputs), JC Richards and Schouten-van Parreren (both cited 20 times), Krashen (cited 18 times), Gairns & Redman (cited 17 times), AD Cohen (cited 16 times) and JB Carroll, Herman, Kucera and MJ McCarthy (each cited 15 times). This list of Highly Cited Influencers shows a remarkable shift from the previous year's data. In the 1992 data set, we identified 10 Very Significant Influencers. Only one of these appears as an Influencer in the 1993 list; the other nine **do** appear in the 1993 data, but with a greatly reduced profile (see Table 15). The most cited influencers for 1993 are listed in Table 16. Of these, only Nation appeared in the 1992 Most Cited Authors list. All these Influencers are familiar from earlier reports, but all of them have significantly increased their presence in the 1993 data. This analysis strongly confirms that 1992 should indeed be treated as an anomalous year and that 1993 is basically a return to normal.

Table 14: *The number of Influencers cited N times in the 1993 data set*

N	**30+**	**29**	**28**	**27**	**26**	**25**	**24**	**23**	**22**	**21**
sources	1	1				1	1	1		1
N	**20**	**19**	**18**	**17**	**16**	**15**	**14**	**13**	**12**	**11**
sources	1		1	2	1	4	2	4	10	4
N	**10**	**9**	**8**	**7**	**6**	**5**	**4**	**3**	**2**	**1**
sources	12	10	13	33	35	50	87	133	386	2065

For the next step in our analysis, we eliminate influencers who are only infrequently cited in the data set, and work in more detail with a reduced set of highly cited authors. By convention, it is customary to work with a set of about 100 highly cited authors, but our analyses of both the 1992 and the 1991 data used a lower figure (78 sources in 1991 and 76 sources in 1992), and in the interests of continuity, it is useful for us to work with a comparable figure from the 1993 data set. The data reported in Table 14 suggests that the closest we can get to this target in the 1993 data is a set of 71 influencers who all recorded at least eight citations in 1993. This inclusion threshold is a substantial increase on the threshold used for our 1992 report (six citations), which reflects the increase in the overall number of outputs that make up the data set. The general features of the 1993 data set are summarised in Table 15, along with the corresponding figures for 1992. Table 16 records the performance of these influencers in the 1992 data set.

Table 15: *The most cited influencers in 1992 and their citations in the 1993 data set*

	Kirsner	Kolers	Lockhart	Smith	Feldman	King	Nation	v Eckardt	Potter	CHen
1992	20	18	17	17	16	16	16	15	14	14
1993	4	8	10	4		8	41	7	7	4

Table 16: *The most cited influencers in 1993 and their citations in the 1992 data set*

	Nation	Laufer	Carter	Anderson	Meara	Nagy	Richards	Schouten	Krashen	Gairns	Redman
1993	41	29	25	24	23	21	20	20	18	17	17
1992	16	11	8	3	12	8	6	6	12	2	2
1991		11	12		15		11		8	9	9

Figure 3 shows a mapping of the 71 authors who are cited at least 8 times in the 1993 data set and identifies cases where these authors are frequently co-cited. In the interests of simplicity only edges that have a weighting of four or more have been included in this mapping

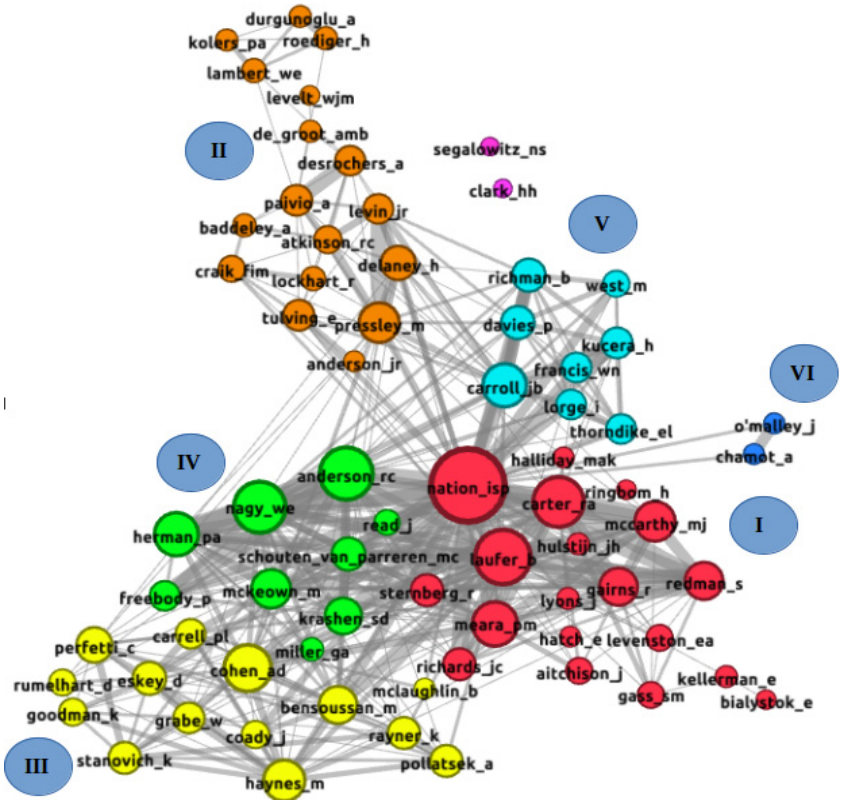

Figure 3: *The co-citations between the 71 most significant influencers in the 1993 data set*

Table 17: *The main features of the 1992 and 1993 eligible data sets*

	1992	1993
Influencers included	76	71
Inclusion Threshold	6	8
New Influencers		40
Lost Influencers	45	

Readers who are familiar with this series of reports will immediately recognise that this mapping looks very different from the 1992 mapping discussed in our last report. The 1992 mapping was completely dominated by a surprisingly large number of psycholinguistic sources, with research in the tradition of applied linguistics reduced to a small cluster on the edge of the main mapping. At the time, we queried whether this was a change of direction, or a temporary blip. The mapping in Figure 3 seems to confirm that the 1992 data was indeed an anomalous blip, and the 1993 mapping looks much more like the 1991 mapping reported in Meara (2023). The 1993 data set sees the return of some familiar themes that we have seen in earlier mappings.

The main feature to note in Figure 3 is the central role of Nation. Nation played only a relatively minor role in the 1992 mapping, but here he has emerged as by far the Most Significant Influence in the field. This rise to prominence is almost entirely due to citations of Nation's seminal text book *Teaching and Learning Vocabulary* (Nation 1990). Other features worth noting in this map are the very strong co-citations between Cluster I and Cluster IV, the very few co-citations between Cluster II and Cluster III, the re-appearance of a word count cluster (Cluster V) and the emergence of a new cluster pointing to strategy research (Cluster VI)

One effect of this return to "normality" is that there is once again a huge turn-over in the list of influencers appearing in the map. Of the 76 names that appear in our mapping of the 1992 data set, only 31 also play a role in the 1993 data set; more than half the names in the 1993 mapping are new (see Table 18).

Table 18: *changes in the composition of the 1992 and 1993 data sets*

Influencers appearing only in the 1992 list (45)

ALTARRIBA BALOTA BEAUVILLAIN BESNER CARPENTER CHADHA CHEN COLLINS COLTHEART CORTESE CURLEY DAVELAAR FELDMAN FORSTER GERARD GRAINGER GROSJEAN HO JAIN KING KIRSNER KROLL LEUNG MACNAMARA MAGISTE MCCLELLAND MEYER NAS NEELY NG PARADIS POTTER REY RUDDY SCARBOROUGH SCHVANEVELDT SCHWANENFLUGEL SEIDENBERG SHARMA SINCLAIR MC SMITH SNODGRASS SO STEWART VON_ECKARDT

Influencers appearing in both lists (31)

ANDERSON ANDERSON ATKINSON BADDELEY BIALYSTOK CARTER DESROCHERS DE_GROOT DURGUNOGLU FRANCIS HERMAN KELLERMAN KOLERS KRASHEN KUCERA LAMBERT LAUFER LOCKHART MCLAUGHLIN MEARA MILLER NAGY NATION PAIVIO PRESSLEY RICHARDS ROEDIGER RUMELHART SCHOUTEN_VAN_PARREREN SEGALOWITZ WEST

Influencers appearing only in the 1993 list (40 : previously unlisted sources shown in **bold)**

AITCHISON BENSOUSSAN **CARRELL** *CARROLL* **CHAMOT CLARK COADY** *COHEN CRAIK* **DAVIES** *DELANEY* **ESKEY FREEBODY** *GAIRNS* **GASS** *GOODMAN* **GRABE HALLIDAY** *HATCH* **HAYNES HULSTIJN LEVELT** *LEVENSTON LEVIN* **LORGE** *LYONS MCCARTHY MCKEOWN* **O'MALLEY** *PERFETTI* **POLLATSEK RAYNER READ** *REDMAN* **RICHMAN** *RINGBOM* **STANOVICH** *STERNBERG* **THORNDIKE** *TULVING*

As usual, the complexity of this map makes it difficult to analyse in detail, and for this reason, a simplified version is provided in Figure 4. This simplified mapping is a spanning tree in which every node in the data set is shown with only its strongest connection to other nodes.

The main structural features to note in this mapping are the central role of Nation, the relatively unstructured nature of Cluster I, the emergence of Laufer, Pressley, Carter and Haynes as secondary hubs, and the marginal position of Cluster VII.

Gephi's formal analysis of this spanning tree in Figure 4 identifies eight thematic clusters in the data.

Cluster I, the largest cluster in the data set (28 members), is predominantly concerned with L2 reading, but it can probably be best seen as comprising five sub-clusters. Thorndike & Lorge, Kucera & Francis, Carroll, Davies & Richman and West are word frequency counts for English. Eskey, Goodman, Coady and Rumelhart are reading specialists. Clark, Lyons and Hatch focus on semantics. Meara, Schouten-van Parreren, Ringbom and Read are loosely linked by an empirical methodology focus.

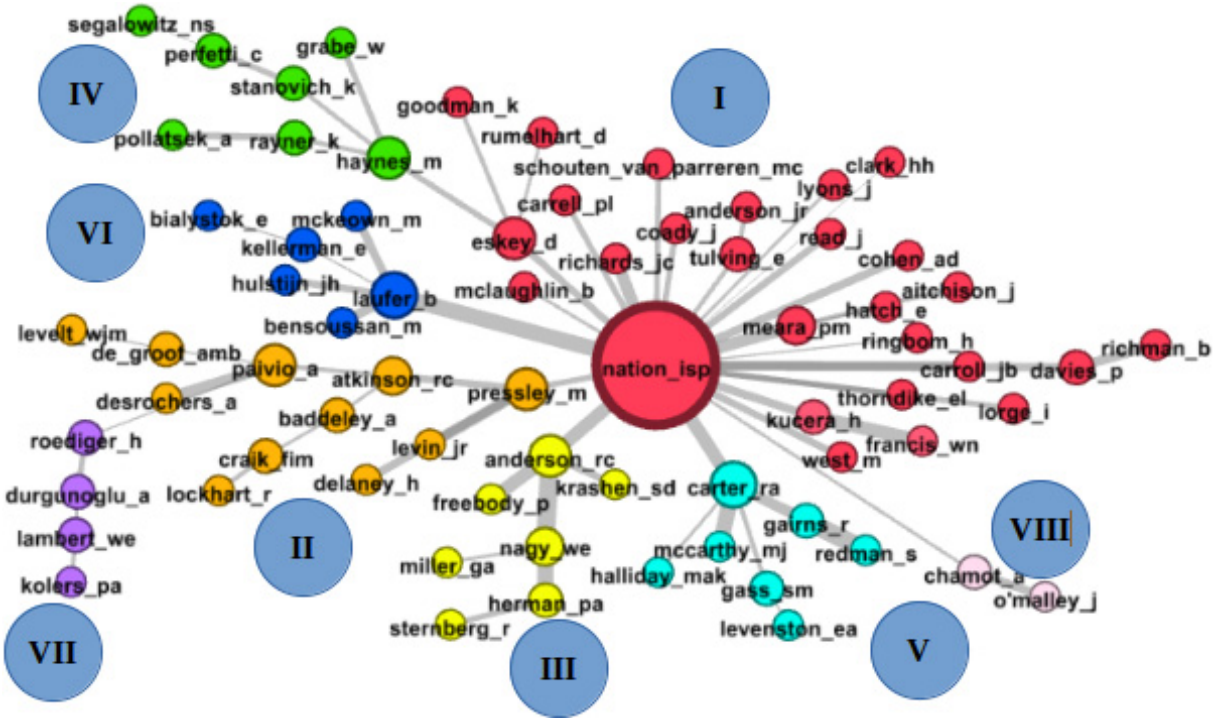

Figure 4: *A spanning tree analysis of the 1993 data set*

Cluster II (11 members), mainly focussed on Pressley, is principally concerned with memory effects in L2 vocabulary acquisition. The prominence of this cluster probably reflects the publication of the conference proceedings edited by Chapelle and Claes.

Cluster III (7 members), focussed on Anderson, is essentially an L1 reading cluster.

Cluster IV (7 members), centred on Haynes, is a set of L1 reading researchers whose work mainly focusses on the psychology of reading. Pollatsek & Rayner are an early example of studies involving eye-movements in reading.

Cluster V (7 members), centred on Carter, is a set of sources whose main focus is the linguistics of vocabulary acquisition.

Cluster VI (6 members), centred on Laufer, is a set of L2 vocabulary researchers, mainly interested in lexical inferencing.

Cluster VII (4 members), centred on Roediger, is a set of North American researchers mainly concerned with bilingual lexicons and memory structures.

Cluster VIII, the smallest cluster in this data set (2 members), signals the importance of learning strategies in this year's data set.

The clusters are summarised in Table 19.

Table 19: *The clusters identified in the 1991 and 1992 data sets*

Cluster	1992	1993
I	Formal models of lexical storage (14)	L2 reading (27)
II	More formal models of lexical storage (13)	Imagery, Mnemonics and L2 vocabulary acquisition (11)
III	Vocabulary uptake and inferencing (13)	L1 Reading (7)
IV	Performance of bilinguals (11)	Psychology of L2 reading (7)
V	Chinese/French (4)	Linguistic approaches to vocabulary (7)
VI	Translation effects in bilinguals (4)	Lexical inferencing and transfer (6)
VII	Imagery and Mnemonics (4)	Bilingual lexicons and memory. (4)
VIII	Memory processes and skilled reading (4)	Strategies in L2 vocabulary learning (2)
IX	Cognate effects (3)	
X	Word Frequency Count (2)	

Table 20: *The strongest co-citation links in the 1991, 1992 and 1993 data sets*

Link Weight	1991	1992	1993
17		Kirsner~Smith	Nagy~Anderson **Gairns~Redman**
16		Lockhart~Smith	
15		King~Smith Feldman~Smith	**Nation~Laufer** **Carter~McCarthy** Nagy~Herman
14		Potter~von Eckardt Kirsner~Kolers	
13		Kirsner~von Eckardt	Nation~Anderson **Kucera~Francis** Nation~Carter Anderson~Freebody
12		Gerard~Scarborough Kirsner~Scarborough Jain~King Feldman~So	**Nation~Meara** Nation~Richards
9	Gairns~Redman		
7	Carter~Meara Carter~McCarthy Carter~Aitchison		
6	Nation~Laufer Nation~Meara Nation~Schouten-van Parreren Raugh~Atkinson Levin~Pressley Kucera~Francis		

The strongest co-citation links in Figure 4 are listed in Table 20, along with the equivalent data for the 1991 and 1992 data sets. The table shows a marked increase in the number of strong co-citations from 1991 to 1993. Strong co-citations in 1993 that also appeared in 1991

are shown in **bold.** The very strong co-citation links joining Gairns & Redman, Nation & Laufer, Carter & McCarthy and Nation & Meara were already present in the 1991 data.

4. Discussion

The simplest account of the changes that have taken place between 1992 and 1993 is that the sudden surge in psycholinguistic research recorded in 1992 was an aberration, and that 1993 represents something of a return to normality. In reality, things are more complicated than this simple picture suggests. In the early 1990s, research ran on a five- or six-year cycle. A significant paper would get published in year X; then it would take another couple of years for readers to react to the original paper and submit follow-up papers in year X+2; these papers would then be rewritten in year X+3 following a lengthy review process; finally, the new version would eventually appear in print in year X+5 or X+6 after another lengthy period waiting in a publication queue. Some journals were notoriously slow in getting accepted articles into print, and a delay of two years was not considered unusual - a stark contrast with the rapid turnaround which characterises more recent research. By way of an example, Umbel, Pearson, Fernández and Oller (1992) - which went on to become a much-cited paper – was cited only twice in 1993, three times in 1994, and once in 1995, and in each of these cases, the author of the citing article was one of the authors of the original paper. It is not entirely surprising, then that the psycholinguistic work cited in 1992 is not frequently cited in 1993. However, we might expect another psycholinguistic ripple to appear in the data for later years.

At the same time, the re-appearance of applied linguistic research in the 1993 map should not be interpreted as the status quo being re-instated. Between 1991 and 1993, we can see a decline in the importance of computational approaches to describing English, and a shift towards lexical inferencing and L2 reading processes. The one feature that maintains a strong presence across all three years concerns the role of mnemonics and imagery in L2 vocabulary learning. Strategies for learning vocabulary emerges as a new theme in the 1993 data set.

Four other features in the 1993 maps are worth highlighting. The first of these features is the dramatic rise of Nation. In 1991 Nation appeared in relatively weak links with Meara and Schouten-van Parreren. In 1992, he does not appear in the list of the strongest co-citation links. In 1993, more than half the strong citation links involve Nation, and he appears as the central focus of the largest cluster in the 1993 map. All the clusters in this map eventually lead to Nation. It is difficult to understate how spectacular this refocussing is.

The second feature worth noting is the emergence of Laufer as a Very Significant Influencer. Her co-citations with Nation were modest in 1991. In 1992, she appears as a weak co-citation with Nation. In 1993 she appears in a very strong co-citation link with Nation, and as the focus of a cluster that contains some of the most important influencers in the vocabulary research so far: Bensoussan, Bialystok, Kellerman, Hulstijn and McKeown. This cluster feels like an important growth point for future years.

The third point worth noting is the importance of Nagy and Anderson in the 1993 maps. The co-citation that links Nagy and Anderson is one of two outstandingly strong links in

Table 20, and the core of Cluster III in Figure 4. This cluster is unusual in that it is strongly linked to Nation, but the members of the cluster are not predominantly L2 researchers. Rather they make up a group of influencers who mainly write about L1 reading.

The fourth point worth noting is the continued importance of Gairns and Redman in 1993. This collaboration - a text first published in 1986 – provided the strongest co-citation link in the 1991 data set. In 1993, it still provides the strongest co-citation in the data set. Sixteen different authors are responsible for the 17 co-citations of Gairns & Redman recorded in Table 20. It is very unusual for a text of this type to be cited so extensively and so consistently.

Although the overall picture for 1993 looks fairly stable, two features suggest that new developments might be in sight on the horizon. Firstly, the 1993 data set includes a substantial number of psychologists working on second language learners. Examples of this are Bahrick, Bahrick, Bahrick and Bahrick (1993), Ellis and Beaton (1993a, 1993b), Service and Craik (1993) and Tinkham (1993). These authors all work on disparate areas of L2 vocabulary acquisition, however, and since they do not cite each other, they do not appear as an identifiable methodological cluster in the 1993 maps. Nonetheless, taken together they look like an alternative take on L2 vocabulary acquisition, and we might predict that this feature will become more prominent in future years.

Secondly, and perhaps more importantly, 1993 sees the appearance of two papers that introduce new formal testing instruments. Paribakht and Wesche (1993) present a self-report tool, the *Vocabulary Knowledge Scale* (VKS), which will become one of the most widely used tools in future research. Read (1993) is the first of a series of studies in which Read outlines his *Word Associates Test*. This tool is designed to test vocabulary depth in a reliable and efficient way, a feature of vocabulary knowledge that has received little attention up until this time. The reason these developments are important is that most previous work has relied on one-off testing materials which vary from paper to paper. This practice makes it almost impossible to construct a coherent narrative out of the increasing number of empirical studies appearing in our data sets. The emergence of "standard" testing tools feels like an important step towards standardisation of the field. (cf. Meara 1993).

5. Conclusion

The 1993 data set has confirmed that the huge shift in focus reported in 1992 seems to be a temporary phenomenon. The psycholinguistics group focussed on Kirsner still appears in the 1989–1993 five-year map, but it does not seem to have developed into a serious, permanent reconfiguration of the field. The split between Applied Linguistic approaches to vocabulary acquisition and more psycholinguistic approaches remains an important characteristic of the research.

The field continues to grow in 1993, and some significant new influencers are appearing in the updated maps. Nation's book *Teaching and Learning Vocabulary* (Nation 1990) is beginning to have an impact on the pattern of citations in 1993, and we can expect this seminal book to be even more influential in coming years. At the same time, some important

influencers from previous years no longer appear in the 1993 map. Corder is perhaps the most important loss, suggesting as it does a decline in the importance of the Department of Applied Linguistics at Edinburgh University. The new focus for L2 research seems to be Wellington, New Zealand.

In summary, 1993 is not a pivotal year for L2 vocabulary research. However, the data suggests that 1993 is not just a return to the status quo either. Nation has become firmly established as the Most Significant Influencer in the field, but overall traditional Applied Linguistic approaches vocabulary seem to becoming less influential than they were in our earlier reports, and new strands of enquiry led by psychologists look as though they might become increasingly influential in future years.

References

Bahrick, H. P., L. E. Bahrick, A. S. Bahrick, and P. E. Bahrick. 1993. Maintenance of foreign language vocabulary and the spacing effect. *Psychological Science* 4: 316–321.

Bastian, M., S. Heymann, and M. Jacomy. 2009. Gephi: An open source software for exploring and manipulating networks. *International AAAI Conference on Weblogs and Social Media*.

Ellis, N. C., and A. Beaton. 1993a. Factors affecting the learning of foreign language vocabulary. *Quarterly Journal of Experimental Psychology* 46A: 533–558.

Ellis, N. C., and A. Beaton. 1993b. Psycholinguistic determinants of foreign language vocabulary. *Language Learning* 43: 559–617.

Lotka, A. J. 1926. The frequency distribution of scientific productivity. *Journal of the Washington Academy of Sciences* 16(12): 317–324.

Meara, P. M. 2023. Shifting sands: A bibliometric analysis of L2 vocabulary research in 1991. *Linguistics Beyond and Within* 9: 112–132.

Meara, P. M. 2024. Laying in a new course? A bibliometric analysis of L2 vocabulary research 1988–92. *Linguistics Beyond and Within* 10: 89–111.

Meara, P. M. no date. VARGA: The Vocabulary Acquisition Research Group Archive. Available at: https://www.lognostics.co.uk/varga/ (accessed 1 October 2025).

Nation, I. S. P. 1990. *Teaching and learning vocabulary*. Rowley, Mass.: Newbury House.

Paribakht, T. S., and M. B. Wesche. 1993. The relationship between reading comprehension and second language development in a comprehension-based ESL program. *TESL Canada Journal* 11, 1: 9–29.

Picoche, J. 1992. *Précis de lexicologie française. L'étude et l'enseignement du vocabulaire*. Paris: Nathan.

Prim, R. C. 1957. Shortest connection networks and some generalisations. *Bell System Technical Journal* 36(6): 1389–1401.

Read, J. 1993. The development of a new measure of L2 vocabulary knowledge. *Language Testing* 10, 3: 355–371.

Service, E. 1993. Phonological and semantic aspects of memory for foreign language. In J. Chapelle and M. T. Claes (eds.), *Memory and memorization in acquiring and learning languages*, 307–318. Louvain-la-Neuve: CLL.

Service, E., and F. I. M. Craik. 1993. Differences between young and older adults in learning a foreign vocabulary. *Journal of Memory and Language* 32: 608–623.

Small, H. 1973. Co-citation in the scientific literature: A new measure of the relationship between two documents. *Journal of the American Society for Information Science* 24: 265–279.

Tinkham, T. N. 1993. The effect of semantic clustering on the learning of second language vocabulary. *System* 21, 3: 371–380.

Umbel, V., B. Z. Pearson, M. C. Fernández, and D. K. Oller. 1993. Measuring bilingual children's receptive vocabulary. *Child Development* 63 (1992): 1012–1020.

Appendix 1

The co-citation methodology

In brief, the analyses focus on the authors cited in the bibliographies of a list of papers published in the relevant time frame, and identify sets of authors who are frequently cited together. These frequently cited authors are deemed to be **Influencers.** Usually, a small number of very strong co-citation clusters emerge from this approach, and we take these clusters to be indications of important research themes in the dataset. The maps developed for this report follow the same format as last year's report, where I used spanning trees (see Appendix 3) based on only the strongest links between the influencers being cited. Readers of this series of reports will realise that the data sets we are describing are becoming increasingly complex as time goes on, and much more difficult to map in a helpful, intuitive way. The spanning tree approach provides a possible solution to this problem.

Appendix 2

Lotka's model

Lotka (1926) suggested that there might be a straightforward relationship between the number of authors who contribute a single paper to a field and the number of authors who make multiple contributions to the field. Suppose, for example, that we have 250 authors who each make a single contribution to a data set, then it would be unusual to find only one author who makes two contributions, and it would likewise be very unusual to find a single author who makes twenty contributions, while no other authors make more than one contribution to the data set. Lotka suggested that the expected relationship could be described as a power law:

$$E_N = T / N^x$$

where T is the total number of authors who contribute a single paper to the data set,
N indicates 2, 3, 4, 5... outputs,

and

E_N is the expected number of authors contributing to N outputs.

In practice, the value of x (the exponent in Lotka's formula) is usually around 2 – that is, a value of 2 for this exponent gives a fair approximation to what happens in real life. So, for a data set in which 250 authors contribute to just one paper in the data set Lotka's model predicts that we can expect 250/22= 63 authors who contribute to two papers in the data set, 250/32 = 28 authors who contribute three papers to the data set, 250/42 = 16 authors making four contributions to the data set, and so on as shown in the table below.

***Table A**: An illustration of Lotka's Law with x = 2 and N_1 = 250*

contributions	**10**	**9**	**8**	**7**	**6**	**5**	**4**	**3**	**2**	**1**
Expected E_N	2	3	4	5	7	10	16	28	63	250

Clearly, this model predicts that the number of papers an active researcher might be expected to produce falls off rather quickly. Empirical tests of what has become known as "Lotka's Law" do seem

to work well. However, the model works best when we are dealing with well-established fields, and very large data sets. The single year data sets that I have discussed in this series of papers are not a close match to Lotka's expectations, but the larger 5-year data sets are generally a better fit to the power law model. In both cases, however, we get a much better fit when the value of N^x is raised above 2. For example, we get the best fit for the 1989–1993 data set when x=2.75. This is lower than the equivalent figure for 1988–92 (x=2.9), though both figures need to be treated with some caution because the data sets are relatively small. Higher values of x seem to be typical of immature, highly volatile fields. Generally speaking, the exponent values we find for the L2 vocabulary research literature are higher than we would normally expect, with a downward trend, and the continued fall in the 1989–93 figure seems to suggest that the field is becoming increasingly "normal".

The data for 1989–93 are displayed below. The table shows the number of authors contributing to N outputs in 1989–93. 578 authors contribute just a single publication in this period. The table also shows how any authors would be expected to make N contributions, given this figure when x=2.

contributions	**20**	**19**	**18**	**17**	**16**	**15**	**14**	**13**	**12**	**11**
Actual 89–93	**2**									**1**
Lotka's model	*1*	*1*	*1*	*2*	*2*	*3*	*3*	*3*	*4*	*5*
contributions	**10**	**9**	**8**	**7**	**6**	**5**	**4**	**3**	**2**	**1**
Actual 89–93	1	1		2	7	9	10	27	87	578
Lotka's model	*6*	*7*	*9*	*12*	*16*	*23*	*36*	*64*	*144*	

It is worth noting here that the 1989–93 data is still disproportionately made up of authors who contribute to a single output, and despite the marked increase in the number of authors contributing to two and three outputs, the actual figures are still about half the of the expected value. Also worth noting are the contributions made by Laufer and Meara. Their twenty contributions each in 1989–93 are best described as heroic.

Appendix 3

Spanning trees

The maps presented in this paper are a simplification of the maps that appeared in the earlier papers in this series. The earlier maps tried to capture the relationships between the Influencers by including any co-citation link which was stronger than a chosen threshold value – for example, we might include any link with a weighting of 8 occurrences or more in the data set. The threshold values were chosen to avoid cluttering up the visuals with very weak connections, but they varied from one report to another, and were essentially arbitrary.

In this paper, I have adopted an alternative solution to this problem, by displaying the data in the form of a spanning tree. In this alternative approach, we start with a list of authors, a list of all the co-citation links between them, ordered by their weight, and an empty map containing no nodes. We then build a map by working through the ordered list of links, and following the steps outlined in an algorithm developed by Prim (Prim: 1957). Starting with the strongest link, we add nodes and edges to the empty map as long as the new edge does not lead to a cycle. That is, if we have a new edge A ~ B, and our tree does not already contain a link (direct or indirect) between node A and node B, then we add the edge A ~ B to the map, adding new nodes as necessary. The map grows in a piecemeal way at

first, adding pairs of strongly connected nodes to the map, but eventually, the algorithm finds a set of links that connects each node to another by its strongest connection.

Acknowledgements

Special thanks to John Read for some very insightful comments on an early draft of this and previous reports. Thanks to Jan Hulstijn who lent me a copy of Chapelle & Claes 1993.

Cite this article as:

Meara, P. (2025). Back to Normality? A Bibliometric Analysis of L2 Vocabulary Research 1989–93. *LingBaW. Linguistics Beyond and Within, 11*, 119–140.