

語彙空間構造を把握する

— 『爾雅』 釋詁・釋言・釋訓の場合 —

谷本 玲大

■ ノードとリンク

- 【ノード (node)】
ネットワークを構成する一つ一つの要素のこと。通信ネットワークではコンピュータやハブ、ルータなど一台一台の通信機器がノードに当たる。ノードとノードを結ぶ線はリンクという。
- 【リンク ; ハイパーリンク (link, hyperlink)】
文書内に埋め込まれた、他の文書や画像などの位置情報。ハイパーリンクを用いて複数の文書、および関連する画像などのオブジェクトを関連付けたシステムをハイパーテキストという。WWW はハイパーテキストの代表例で、Web ブラウザで文書を表示し、リンクのある場所を、マウスでクリックすると、関連づけられたリンク先にジャンプするようになっている。
(IT用語辞典 e-Words <http://e-words.jp/>)

国文学研究者にとってはノードという言葉はあまり馴染みがないかもしれない。

最も代表的な WWW の HTTP プロトコルで送受信され、グラフィカルブラウザで表示・閲覧される HTML ファイルによる WebPage を例に言えば、各 HTML ファイルがノード、それらを結んでいるのがリンクということになる。

国語・国文学研究は、その扱うコンテンツの特性から、ノード同士のリンク情報を重視せざるを得ない。

例えば国文学に於いては、王朝貴族の人生を左右した血脈や主家との関係、古今伝授・琴・手習い等の師の系譜、僧侶の法系など。

国語学にあつては、語彙の意味範囲・文法体系に於ける品詞論など。

いずれも相互関係性の中でこそ、当該ノード（人物なり語なり品詞なり）の位置づけが可能になる。

本稿では、この抽象化概念を援用して語彙空間について考えてみたい。

ここで言う語彙とは、一定のボキャブラリの集合体であるところの語彙（例えばある国語辞典に収録される範囲の語彙など）を言い、語彙空間とは、それら各単語相互関係を為して織りなす結束線の繋がり方をイメージしている。

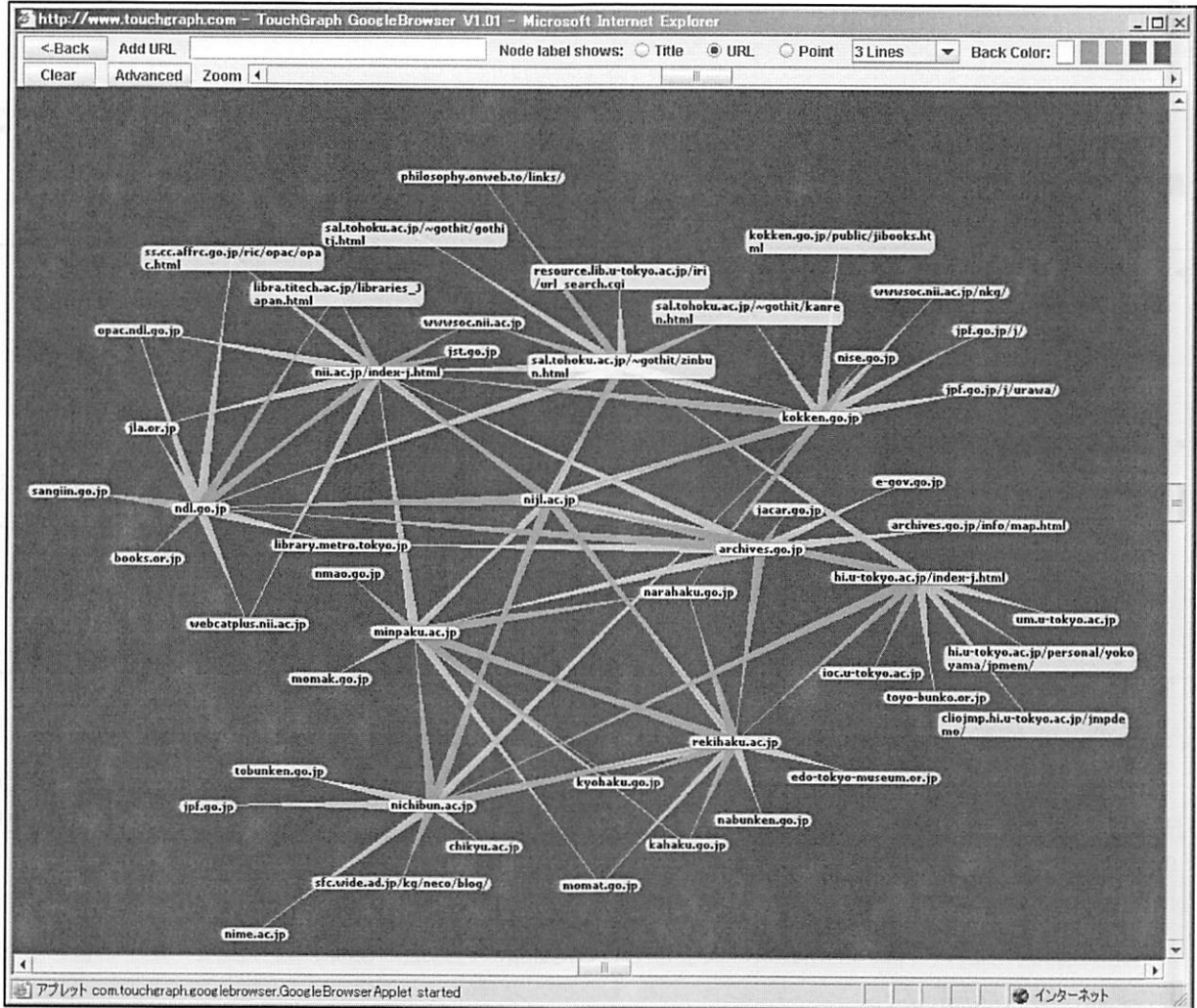
■ リンクの視覚化

以下は、サーチエンジン Google によって国文学研究資料館の WebPage (<http://www.nijl.ac.jp/>) と類似・関連するものとして分類されているサイト群を視覚化したものである。

この視覚化は JavaApplet ツール “TouchGraph GoogleBrowser V1.01”¹ を実行して行った。

¹ <http://www.touchgraph.com/TGGoogleBrowser.html>

“TouchGraph GoogleBrowser” は、恐らく “related:www.nijl.ac.jp/” として Google に検索クエリを送り、その結果として上位に出てきたサイトについても再帰的に “related:” の検索指令を出しているものと推測されるが、それら関連情報をどのように取り扱って類似・関連と見なしているかについては今は敢えて立ち入らず、その結果を見ることにしよう。



結果は上図の通りである。

国文学研究資料館と同じ「大学共同利用機関法人間文化研究機構」に属する国際日本文化研究センター (<http://www.nichibun.ac.jp/>)、国立民族学博物館 (<http://www.minpaku.ac.jp/>)、国立歴史民俗博物館 (<http://www.rekihaku.ac.jp/>) の他、国文学研究資料館と組織機能的に類縁関係にある国立公文書館 (<http://www.archives.go.jp/>)、国立国会図書館 (<http://www.ndl.go.jp/>)、国立国語研究所 (<http://www.kokken.go.jp/>)、国立情報学研究所 (<http://www.nii.ac.jp/index-j.html>)、東京大学史料編纂所 (<http://www.hi-u-tokyo.ac.jp/index-j.html>) などの機関がそれぞれ核になりつつリストアップされていること、また、学術系サイトの網羅的リンク集として著名な後藤斉氏（東北大学大学院文学研究科言語学研究室）による国内人文系研究機関 WWW ページリスト (<http://www.sal.tohoku.ac.jp/~gothit/zinbun.html>) が核の一端を担っていることが見て取れる。

このように、Web 空間内部で各サイト同士の占める位置・独自性・色彩付けは、リストされた機関同士の結束線の繋がり方によって支えられていると言え得る。

本稿で後述する語彙空間を考える際に、この点は重要であるので強調しておく。

さて、TouchGraph GoogleBrowser の出力結果が、ここ 10 年ほどにわたって我々が親しんできた WWW ページの検索エンジンとどのように違うのかと言えば、主要なポイントは一点だけである。

それは、各 WebPage をノードとし、「それら同士のリンクがどのように結束しているのか」をも視覚化できたことによって、それら各ノードが存在する Web 空間の全体を視覚的に把握することができるようになったという点である。

尤も、先に敢えて不問とすることを断ったように、今の例で挙げた「結束線の繋がり方」は、それを直接的なリンク関係だけで計算する場合や、各サイト内部のテキストを統計的に処置して類義語の使用率で重み付け処理する場合など、様々な計量の仕方がある。

当然、その様々な計量方式のどれを採用するかで結果が異なってくるのであって、視覚表現の妥当性の検証についてもなお慎重を要しよう。

また、「どちらのノードからどちらのノードに対する結束なのか」という結束の向きの指示・視覚化にまでは至っていない点もすぐに改善要求として思い浮かぶ所であろう。

情報科学技術分野では、これらの問題点を十分に把握した上で類似同士をまとめるクラスタリング作業やその理論、また、その結果を図として表示する可視化ツールの開発などの研究が進められており、一部では他の研究に資する実用ツールとしての活用も進められている。

例えば、細胞内での遺伝子制御ネットワークや代謝ネットワークのような大量の行列データのネットワーク構造を分析する操作などで利用されているようである。

視覚化ツールの様々については、増井俊之氏²のサイトにある「視覚化関連文献」や、佐藤克己氏³のサイト、小池英樹氏⁴の解説などを参照のこと。

■語彙空間把握の考え方—『爾雅』の場合—

本稿では、視覚化の達成を最終目標とはしない。

主たる目的は、語彙空間内部に於いて、各語の存立を支える結束線（リンク）情報に支えられる意味ネットワークの抽出をすることである。

視覚化システムへの実装は、その次の段階で必要になるであろうが、今回はそれ以前の基礎検討について報告を行う。

実は当初、本稿では「国語辞典」の見出し語と、その語義説明との関係を解析することを企図し、下作業として EPWING 形式の『広辞苑 第四版』、『大辞林 第二版』のテキスト抽出を合法的な手段で行なって、分析を試みた。

それによって、例えば、「見出し語としては表層に現れていないが、実は説明語として様々な意味範疇の中で用いられる融通無碍な単語」が見つかる可能性に期待を寄せたのである。

しかし、説明文にある「言い換え」には規則性がなく、長大な名詞句を形成するもの、「～こと。」「また、～。～とも。」など、単純に言い換えを行っていないために「～」で述べられるものにほぼ該当する見出し語があっても同一ノードとして判定できず、ためにリンクの認定が行えない場合が多くあるなど、データ分割方針に客観性を保障する方策が見出せず、解析を断念した。

² <http://pitecan.com> の <http://pitecan.com/bib/Visualization.html> や <http://pitecan.com/presentations/D EWS98/index.html> など。

³ <http://apollo.u-gakugei.ac.jp/~yoshiki/research/survey-iv.php>

⁴ 小池英樹氏「ビジュアルライゼーション」(『bit 別冊 ビジュアルインタフェース—ポスト GUI を目指して—』平川・安村編、pp.24-44、共立出版、1996、<http://www.vogue.is.uec.ac.jp/~koike/bit/bit.html>)などを参照。

そこで今回は、中国最古の字書である『爾雅』を用いることとする。

【爾雅】じが

十九篇。〔漢書芸文志〕に〔爾雅〕三卷二十篇とするが、いま十九篇を存する。撰者不詳。最古の字書。釈詁・釈言は字義、釈訓は重言、釈親・釈宮以下その事類をいう。〔詩〕の訓義をとることが最も多く、古い訓注を集めたものであろう。成立について内藤湖南の〔爾雅の研究〕がある。晋の郭璞の〔注〕、宋の邢昺の〔疏〕は〔十三經注疏〕の一。清の邵晋涵の〔爾雅正義〕二十卷、郝懿行の〔爾雅義疏〕十九卷がある。(CD-ROM版『字通』付録 白川静著 平凡社 2003/07)

『爾雅』の構成は釋詁、釋言、釋訓、釋親、釋宮、釋器、釋樂、釋天、釋地、釋丘、釋山、釋水、釋草、釋木、釋蟲、釋魚、釋鳥、釋獸、釋畜である。

今回は、この内、比較的単字が多く、語としての抽象度が高い字を集めている先頭三篇を検討対象とした。

底本には、尊經閣文庫旧蔵、東京大学東洋文化研究所蔵、崇禎中古虞毛氏汲古閣刊十三經註疏本『爾雅註疏』十一巻を用いた。

釋詁、釋言、釋訓の親見出しは617個で、その構成は大凡、以下のようになっている。

太字については補足説明がある。それについては後述する。

仮番号	分類	見出し	説明
1	釋詁	始	初、哉、首、基、肇、祖元、胎、俶、落、權、輿
2	釋詁	君	林、烝、天、帝、皇、王、后、辟、公、侯
3	釋詁	大	弘、廓、宏、溥、介、純、夏、幬、彫、墳、嘏、丕、弈、洪、誕、戎、駿、假、京、碩、濯、訐、宇、穹、王、路、淫、甫、景、廢、壯、豕、簡、筭、販、咍、將、業、席
4	釋詁	有	幬、彫
5	釋詁	至	迄、臻、極、到、赴、來、弔、艘、格、戾、懷、摧、詹
6	釋詁	往	如、適、之、嫁、徂、逝
7	釋詁	賜	賚、貢、錫、畀、予、貺
8	釋詁	善	儀、若、祥、淑、鮮、省、臧、嘉、令、類、緝、穀、攻、穀、介、徽
9	釋詁	敘	舒、業、順
10	釋詁	緒	舒、業、順、敘
11	釋詁	樂	怡、懌、悅、欣、衍、喜、愉、豫、愷、康、妯、般
12	釋詁	服	悅、懌、愉、釋、賓、協
13	釋詁	自	通、遵、率、循、由、從
14	釋詁	循	通、遵、率
15	釋詁	謀	靖、惟、漠、圖、詢、度、咨、諏、究、如、慮、謨、猷、肇、基、訪
16	釋詁	常	典、彝、法、則、刑、範、矩、庸、恆、律、戛、職、秩
17	釋詁	法	柯、憲、刑、範、辟、律、矩、則
18	釋詁	罪	辜、辟、戾
19	釋詁	壽	黃髮、鬣齒、鮐背、耆老
20	釋詁	信	允、孚、寘、展、謀、誠、亮、詢
21	釋詁	誠	展、謀、允、慎、寘

22	釋詁	戲謔	謔浪笑敖
23	釋詁	曰	粵、于、爰
24	釋詁	于	爰、粵
25	釋詁	於	爰、粵、于、那、都、繇
26	釋詁	合	故、郤、盍、翕、仇、偶、妃、匹、會
27	釋詁	匹	仇、讎、敵、妃、知、儀
28	釋詁	對	妃、合、會
29	釋詁	媿	妃
30	釋詁	繼	紹、胤、嗣、續、纂、綏、績、武、係
31	釋詁	靜	忒、謚、溢、蟄、慎谿、謚、顛、頤、密、寧
32	釋詁	落	隕、碩、湮、下降、墜、擗、蓋
33	釋詁	告	命、令、禧、眇、祈、請、謁、訊、誥
34	釋詁	遠	永、悠、迴、遑、遐、邊、闊
35	釋詁	遐	永、悠、迴、遠
36	釋詁	毀	虧、壞、圯、圯
37	釋詁	陳	矢、雉、引、延、順、薦、劉、繹、尸、旅
38	釋詁	主	尸、職
39	釋詁	案	尸
40	釋詁	官	案、寮
41	釋詁	事	績、緒、采、業、服、宜、貫、公
42	釋詁	長	永、業、引、延、融、駿
43	釋詁	高	喬、嵩、崇
44	釋詁	充	崇
45	釋詁	勝	犯、奢、果、毅、剋、捷、功、肩、堪
46	釋詁	克	勝、肩、戡、劉、殺
47	釋詁	殺	劉、獮、斬、刺
48	釋詁	勉	𡗗、𡗗、孟、敦、勗、劔、茂、劭、勗
49	釋詁	強	驚、務、昏、警
50	釋詁	我	卬、吾、台、予、朕、身、甫、余、言
51	釋詁	身	朕、余、躬
52	釋詁	予	台、朕、賚、昇、卜、陽
53	釋詁	進	肅、延、誘、薦、餞、晉、寅、蓋
54	釋詁	進	羞、餞、迪、烝
55	釋詁	導	詔、亮、左、右、相
56	釋詁	勳	詔、相、導、左、右、助
57	釋詁	右	亮、介、尚
58	釋詁	亮	左、右
59	釋詁	光	緝、熙、烈、顯、昭、皓、頰

上記のうち、太字部分を用いて、若干の補足説明をする。

仮番号	分類	見出し	説明
9	釋詁	敘	舒、業、順
10	釋詁	緒	舒、業、順、敘

見出し「叙」に対して単字を列挙しているが、『爾雅』に於いては、これは「舒也。業也。順也。」と示しているのと同じである。

そして次の見出し「緒」に於いて「舒、業、順」と挙げ、更に「叙」も示している。

すなわち、

$$\left\{ \begin{array}{l} \text{「叙」} = \text{「舒」} = \text{「業」} = \text{「順」} \\ \text{「緒」} = \text{「舒」} = \text{「業」} = \text{「順」} = \text{「叙」} \end{array} \right.$$

の二系列の同義字リストが掲出されていることとなり、結果としては「叙」＝「舒」＝「業」＝「順」＝「緒」と、全ての単字が等価の関係を作っていることとなる。

ここで当初から「叙、舒業順緒」と示されていないのは、先引した白川氏解説（内藤湖南説を踏襲）に「古い訓注を集めたものであろう。」とされるのを参考とするより他なく、今は、典拠となった訓注の出た順に追補列挙して行った過程での未整理であろうと見ておく。

このようなものは以下の例にもある。

23	釋詁	曰	粤、于、爰
24	釋詁	于	爰、粤
25	釋詁	於	爰、粤、于、那、都、繇

「曰」は助詞。訓ずるならば「ここに、に、ここにおいて、これ」となる。

以下も同様で、訓ずるならば「とおい、はるか」となる。

34	釋詁	遠	永、悠、迥、遑、遐、邈、闊
35	釋詁	遐	永、悠、迥、遠

更に、別見出しとして以下のような例がある。

42	釋詁	長	永、業、引、延、融、駿
----	----	---	-------------

「長」は「ながい」の他に「トホシ」の訓がある。「融」は「融遠」の熟語があつて「はるかに長い」。「駿」は聲符「夂」に高大、峻高の義があるかと思われる。言うまでもなく訓は日本語への割り当てであつて、その発想・意味範疇のズレが釋詁内部で存在するの否かは訓で見ても判断はつかない。

しかし、このように同義字であろうと見るわけには決して行かない例もある。以下の例、

50	釋詁	我	卬、吾、台、予、朕、身、甫、余、言
51	釋詁	身	朕、余、躬
52	釋詁	予	台、朕、賚、畀、卜、陽

の内、50と51は基本的に「われ、みづから、自分自身」の意である。

しかし、52の内、「賚」「畀」「卜」は「与える」の意であつて、「われ、みづから、自分自身」の意ではない。されば、以下の2つのグループに分かれねばならぬこととなる。

$$\left\{ \begin{array}{l} \text{「我」} = \text{「身」} = \text{「予」} = \text{「台」} = \text{「朕」} \\ \text{「予」} = \text{「賚」} = \text{「畀」} = \text{「卜」} \end{array} \right.$$

以上、『爾雅』を使用する際、本来ならば注意すべき面があることを述べた。

但し、以下の作業ではこれらは機械的に同一視して処理することとし、その上で、『爾雅』内部での「見出し」と「説明中の各語」とをノードとして扱った場合、どのようなリンクが結束されることになるのか

を見てゆきたい。

■『爾雅』をソシオグラムに適用する

上述のようにして整理した『爾雅』データを分析する際に採用するのは sociometry (ソシオメトリー) である。

【ソシオメトリー (sociometry)】

集団内での人間関係における構造上の特性を、成員間の牽引や反発の力学的関係としてとらえ、その量化測定を行う方法と理論。アメリカの心理学者モレノとその学派によって体系化された。社会測定法。→ソシオグラム

【ソシオグラム (sociogram)】

特定の小集団における人間関係のつくる構造を明らかにするため、成員相互間の選択 (牽引)・拒否 (反発) 関係を図表化したもの。ソシオメトリーで使用される。

『大辞林』第二版 三省堂

ソシオメトリック=テストは、小学校の学級運営などで実用的に使用されるケースが多かったようである。実際、稿者が小学生 (三、四年生) の頃、このソシオメトリック=テストは複数回行われた記憶がある。それは以下のようなアンケート調査であった。

- ・ 席替えをしたら、同じグループになりたいと思う人を3人書いてください。
 - ・ できれば同じグループにはなりたくないと思う人を3人書いてください。
- (答えたくないときは、無理に答えなくてもかまいません。)

このようなアンケートは、各学期の中盤に差し掛かると配布され、翌週にクラス内での班替えが行われた。

(今になって思えば、恐らく、) 担任教師は生徒同士の

- ・ 選択のベクトル (一方通行、双方向) の抽出
- ・ 排斥のベクトル (一方通行、双方向) の抽出
- ・ 非選択による孤立者の抽出

…といった作業を行い、子供であるが故になおさら生々しい人間関係の網目をかれこれ勘案して、排斥を解きほぐしたり、中立的立場にある生徒を媒介として新たな人間関係の切っ掛けを紡ごうと工夫するなどして、班構成の案を練ったのであろう。

但し、秋田県内の小学校で教頭を務める佐々木彰氏の WebPage にある解説「ソシオメトリーをもう一度」によると、近年は、人権問題に発展することを避けて、このようなことは行われなくなってきたようである。

以下、佐々木彰氏「ソシオメトリーをもう一度」にある解説を引く。

具体的には、集団の中で各成員が他のどの成員を選択するか (好きか)、どの成員を排斥するか (嫌い) を調査し、その相互関係をまとめる。

一般的には、選択者・排斥者ともに3名程度の名前をあげさせる。

この結果を、図示したものをソシオグラム、表で示したものをソシオマトリクスという。

下の表は、ソシオマトリクスの例である。

この表を縦に見ていけば、その子が誰を選択し、誰を排斥しているかがわかる。例えば2番のB助の場

合は、A男・C造・E彦を選択し、G也・I彦・J次を排斥している。

B助の場合、A男・C造からは選択されているので、相互選択を示す「◎」のマークがついている。また、J次とは相互に排斥しているため、相互排斥を示す「キ」マークになっている。

横に見ていった被選択・被排斥の状況によって、多くの子から選択されている者には「スター」、多くの排斥を受けている者には「排斥児」などの注釈がついている。

赤線で囲んでいるのは、相互選択による結びつきが強い、いわゆる「仲良しグループ」である。その中に入っていない子は、選択・排斥の状況によって、周辺児・孤立児と判断される。

佐々木彰氏「ソシオメトリーをもう一度」(<http://www2.chokai.ne.jp/~assoonas/UC331.HTML>)

		1	2	3	4	5	6	7	8	9	10	被 選 数	被 排 数	相 互 選 斥	相 互 排 斥
	A 男	B 助	C 造	D 夫	E 彦	F 吉	G 也	H 郎	I 弘	J 次					
1	A男	◎	◎	◎	○	キ		○				6	1	3	1
2	B助	◎	◎	×	×		○	○	○	キ		5	3	2	1
3	C造	◎	◎	○			○					4	0	2	0
4	D夫	◎			×			○	○			3	1	1	0
5	E彦		○	○		◎	◎		○	○		6	0	2	0
6	F吉	キ	×	○	◎	◎	×	キ	キ			2	5	1	3
7	G也		×		◎	○	×					2	2	1	0
8	H郎	×		×		○	×	○	○			3	3	0	0
9	I弘		×	×		キ			キ			0	4	0	2
10	J次	×	キ	×	×	キ	×	×	キ			0	8	0	3
	選択数	3	3	3	3	3	3	3	3	3					
	排斥数	3	3	3	2	3	3	2	3	2	3				

この方法を『爾雅』について行った場合、どのようになるか。

Access で計量し、縦列の見出し語と、横列の見出し語が相互参照している場合を示す。

57	釋詁	右	亮、介、尚
58	釋詁	亮	左、右

が繋がる事がわかる。

title_word	右	遠	勤	告	請	多	撫	暮	亮	勞	遐	厭
右									◎			
遠											◎	
勤										◎		
告					◎							
請				◎								
多												◎
撫							◎					
暮								◎				
亮	◎											
勞			◎									
遐		◎										
厭						◎						

実際の所、『爾雅』は並列列挙であるので、「見出し語」を認定するのは正しくない。

今、便宜的に、このように扱ったのである。

なお、この便宜的な「見出し語」と、それ以外の語とを区別せずにどちらかから参照している場合のメトリックを作成すると、縦横 180 マス程度の一覧表ができた。

これは俯瞰性の面で、やや見通しが悪いので、今、煩瑣を避けてこれは掲示しないが、この基礎データを用いて、XML などタグ付けを行い、リンク表示の生成を行うビジュアルライゼーション・ソフトウェア

ア⁵にかければ、語彙空間内部に於いて、各語彙の存立を支える結束線（リンク）情報に支えられる意味ネットワークを視覚化することができよう。

■『爾雅』をインパクト・ファクターで計量する

Impact Factor は、ある雑誌が1論文あたり平均で何回引用されているかを算出し、学界内での、「その雑誌の重要度」を示す指標⁶で、主に理科系分野の業績評価では重要な観点として活用されている。

被引用数が多いほど、その論文は有用である可能性が高く、そのような論文を多く載せている雑誌は、学界への裨益も大きいであろう…という観点に立って計量される。

但し、この数値はあくまでも「その雑誌の重要度」を示すものである。言い換えれば掲載した個別論文の被引用数を根拠にして「雑誌の格」が決定されるのである。

その、格が高い雑誌に自己の論文が載ったということと、その個別の論文の質が高いこととは別問題である点については、インパクト・ファクターの創案者である Eugene Garfield 自身も述べている。

極論であるが、多くの論者が批判的言辞とともに、「全く内容的に取るに足らないダメ論文の代表例」として挙げて引用した場合でも、ポイント数は上がるのである。

また、自己が関係する雑誌のポイントを上げようと意図的な引用を勧める編集委員もいるという。学界の構造的な問題と言えようか。

このように、雑誌のインパクト・ファクターを、個別の論文について敷衍適用して捉える考え方は“誤った使い方”であると指摘されていることを述べたが、実際のところは、我々の身の回りにこのような素朴な考え方の方が根強いと思われる。

例えば、『勅撰作者部類』の編纂意図も、これに類する素朴な思考であろう。

○計算式の例：

$$IF = (\text{万葉集全歌中、人麻呂詠の歌数}) / (\text{万葉集全歌数})$$

$$IF = (\text{勅撰集全歌中、貫之詠の歌数}) / (\text{勅撰集全歌数})$$

このような式を適宜に案出して比較すれば、認定基準をどのようにするかは一考の余地があるにせよ、興味をそそる数値を得られよう。

他に、以下のような計量比較も我々にとっては意味を持つであろう。

○御子左家と六条家の歌壇への影響力の比較

$$\text{御子左家の IF} = (\text{勅撰集全歌中、御子左家歌人の詠歌数}) / (\text{勅撰集全歌数})$$

VS

$$\text{六条家の IF} = (\text{勅撰集全歌中、六条家歌人の詠歌数}) / (\text{勅撰集全歌数})$$

○人麻呂と貫之とどちらの影響力が実作に於いて強いのか？

$$\text{人麻呂 IF} = (\text{新編国歌大観全歌中、人麻呂詠の本歌取り数}) - (\text{新編国歌大観全歌中、人麻呂詠の数}) / (\text{新編国歌大観全歌数})$$

⁵ 例えば graphviz や KeyGraph のような、ばねモデルによるノードの最適配置を適用する視覚化ツールなど。

⁶ 窪田輝蔵氏『科学を計る—ガーフィールドとインパクト・ファクター』（インターメディカル、1996/10）、山崎茂明著『インパクトファクターを解き明かす』（情報科学技術協会、2004/03）、などを参照。通常、 $IF = (\text{ある論文誌に掲載された全ての論文が引用された回数}) / (\text{その論文誌に掲載された論文数})$ で表され、IF 値の実際の計算では、直前の2年間のデータをもとに算出している。例えば、2006年のIF値は、2005年と2004年に掲載された論文が2006年に引用された回数を、同じ2年間の掲載論文数で割った値になる。

VS

貫之 IF = (新編国歌大観全歌中、貫之詠の本歌取り数) - (新編国歌大観全歌中、貫之詠の数)
 / (新編国歌大観全歌数)

以下、このような「《疑似的》インパクト・ファクター」の観点に立って、『爾雅』冒頭三篇のデータを計量する。

まず、「見出し」に出ているある文字自身が、「他の箇所の“見出し”」または「見出しとしては一致しない語の“説明”」のいずれかとして出現したものを抽出する。

その数は617であった。今、その上位を表にして示す。

仮番号	見出し	他の説明との一致数	他の見出しとの一致 (自身を引いた後の数)	単純加算による IF
233	為	5	4	9
☆554	烝	6	0	6
105	業	5	0	5
213	戾	4	1	5
4	有	4	1	5
☆472	眇	3	2	5
378	相	3	2	5
288	慄	2	3	5
448	既	1	4	5
528	勞	2	2	4
239	延	4	0	4
523	安	3	1	4
95	遇	3	1	4
322	會	2	2	4
218	逮	3	1	4
57	右	3	1	4
230	齊	3	1	4
340	邊	2	2	4
32	落	2	2	4
102	微	2	2	4
124	作	1	3	4

次に、「見出し以外の“説明語”として出ているある文字自身が、「他の箇所の“見出し”」または「見出しは一致しない語の“説明語”」のいずれかとして出現したものを計量する。

その数は1084であった。今、その上位を表にして示す。

仮番号	非見出し説明	見出しとの一致数 (自身は含まない)	他の説明との一致 (自身を引いた後の数)	単純加算による IF
1078	𪛗	0	7	7
1079	𪛘	0	7	7
1080	𪛙	0	7	7

1081	梟	0	7	7
1083	跗	0	7	7
1084	翹	0	7	7
☆665	烝	1	5	6
109	業	1	4	5
444	明	3	2	5
464	猷	0	5	5
723	肅	0	5	5
8	安	2	2	4
☆32	延	1	3	4
88	休	0	4	4
285	績	0	4	4
297	相	2	2	4
303	卒	0	4	4
833	齊	2	2	4
922	戾	1	3	4

上のようになる。出現数 617、1084 共に、10 に満たない数値を割るに適さないため、今、そのままの数値で見ることとする。

○「眡」の例。

33	釋詁	告	命、令、禧、眡、祈、請、謁、訊、誥
202	釋言	致	眡、底
393	釋言	殄	眡
472	釋言	眡	障

○但し、「烝」のデータのよう、釋訓で数値を稼いでいるものもある点に注意。

2	釋詁	君	林、烝、天、帝、皇、王、后、辟、公、侯
54	釋詁	進	羞、餞、迪、烝
73	釋詁	眾	黎、庶、烝、多、醜、師、旅
84	釋詁	祭	禋、祀、祠、烝、嘗、禴
377	釋言	塵	烝
518	釋訓	作	烝烝遂遂
554	釋訓	烝	煇煇

○「延」の例。

37	釋詁	陳	矢、雉、引、延、順、薦、劉、繹、尸、旅
42	釋詁	長	永、業、引、延、融、駿
53	釋詁	進	肅、延、誘、薦、餞、晉、寅、蓋
101	釋詁	間	孔、魄、哉、延、虛、無、之、言
239	釋言	延	賈

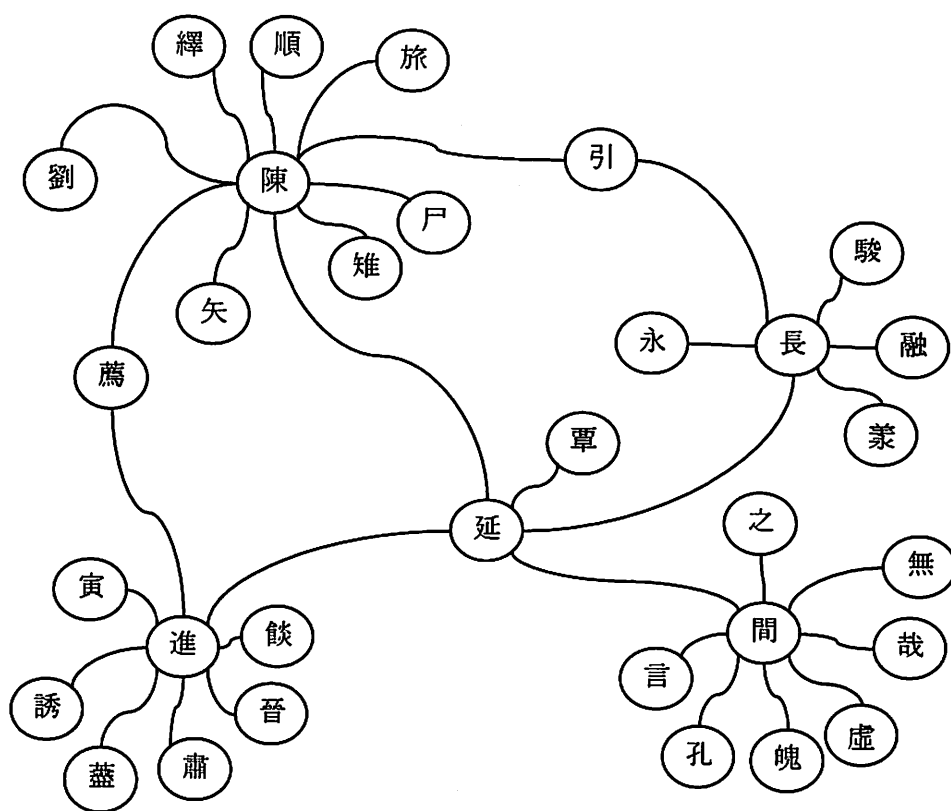
ここで上位に来た文字が、『爾雅』内部で、「使用範囲の広い、融通無碍なる語」として機能している。本稿冒頭で掲示したような視覚化を行った際には、ノードとして立てられた語から伸びているリンク情報の結束線が、取り立てて多くなるはずである。

いわば、核ノードとしての地位を確保することがわかる。

今、仮に手動描写で「延」について図示すると、以下のようになろう。

但し、この図は、本来、「陳」、「矢」、「雉」、「引」、「延」、「順」、「薦」、「劉」、「繹」、「尸」、「旅」などの全てを同等の結束線で結ぶべきだが、図があまりに煩雑になるため、見出しで代表せしめた場合の図として作成したものである。

「陳」、「長」、「進」、「間」がそれぞれのグループを作り、それらが「延」によって結ばれることがわかる。図示によって、「引」も同様の役割を果たしていることも判明する。



以上、『爾雅』の語彙空間を把握するに資する内部構造の検証ができた。

■今後の課題

今後は、オブジェクトとオブジェクトを結んでいるノードを結ぶリンクの結束線を鳥瞰できるようにする視覚化システムの検討、そのリンクが保持しているベクトルが、一方向なのか、双方向なのか、一方向の場合にそれは正のベクトル/負のベクトル（例えば同義・類義/対義など）の何れなのか…なども区別したビジュアライゼーション（視覚化）の実装を検討する必要がある。