計量漢文学への道

160121
島野達雄

1. 漢文との出会い

人はとしをとると，子供のころに帰ってゆくという．わたしもこの十年ほど徐々に中学校小学校のころに意識がさかのぼることが多くなった．漢文とわたしとの出会いをこの機会に書いておこうと思う．

明治 40 年生まれのわたしの父親は，姫路師範学校を卒業し，昭和の初めごろ尋常小学校の教員になった．そのあと，当時の文部省検定試験，略して文検（もんけん，と言っていた）で旧制中学の「国語」ついで「漢文」の教員資格を苦労してとった．家庭では異常なほど子供たちに厳しかった父親ではあったが，「文検の最終面接で東大に入った」という冗談を一度だけ聞いたことがある．家の書棚には『大言海』や『日本文学大辞典』，山田孝雄の『日本文法論』とならんで簡野道明のぶあつい漢文の辞典があった．最初の数ページだけ返り点をほどこした白文の『十八史略』や国訳漢文大成の『論語』もあった．『論語集注（しっちゅう)』もあったと思う．

父が 40 歳すぎのときの 4 人兄弟の末っ子であったわたしは，父親がこれらの本をひろげている姿を見たことがない．文検のあと，これらの受験参考書は書棚をかざっているだけだった．小学校の高学年から中学校にかけて，わたしは家族の誰にも言わずに，こっそりとこれらの本を順にひろげていった記憶がある．いったい何のために簡野道明の本をひもといたのか，今となってはおぼえていない．

2. 受験漢文と漢文の違い

はじめに「漢文」とは何かを論じておこう．

最初に注意しておかねばならないのは，高校で習う，つまりセンター試験などの大学入試に出題される「受験漢文」と，林羅山や荻生徂徠が書いた「漢文」とのあいだには，釈迦に説法かもしれないが，「受験数学」と「数学」と同じぐらいのへだたりがある，という点である．

一例をあげよう．

戦前の中学（旧制中学）の主要 4 教科は，国漢英数つまり国語・漢文・英語・数学であった．10 年ほど前の近畿和算ゼミナールには，キラ星のごとく大正末や昭和ヒトケタ生まれの先生方が出席しておられたが，旧制中学で漢文を学んだはずの先生方の誰一人として，「竪点（たててん)」という言葉をご存知なかった．というのも，戦前からすでに漢文の授業は，「受験漢文」の授業になっていたからである．

かくいう私は，貝原益軒の『点例』を読んで初めて「竪点」という言葉を知った．日本国語大辞典には「たててん」が掲載されている．（竪点の使い方については，次節で紹介する．）

このように，「受験漢文」と「漢文」には，使っている「用語・記号」の違いがある．むろん戦後の新字体ではなく，「漢文」は旧字体の漢字（旧漢字）で書かれている．

正直に言うと，いくら「受験漢文」に長（た）けていても江戸時代の「漢文」は読めないと思う．私もそうとう苦労した．国語でも数学でも教える側にまわるとすぐに気づくが，試験に出す問題は，おおむね「答えが一通り」になるように作る．採点をスピードアップするという教師の都合があるから．

わたしの父親が学んだ漢文は文検受験のための「受験漢文」だった．

この原稿で述べる「漢文学」とは，「受験漢文」ではない「漢文」を研究する学問・科学であるとご理解いただきたい．

3. 漢文を読むために

答えが一通りでない例として，つまり受験漢文では教えない話題として，「有朋自遠方来」「然而」「云爾」の読み方，竪点の使い方，四声の記号，訓読における品詞など，いくつかの事例を紹介しよう．

(1)「有朋自遠方来」の読み方

最初に語順，言いかえると漢字の並び替えの問題を取りあげる．

『論語』学而篇の「有朋自遠方来」を，林羅山や伊藤仁斎は「朋遠方より来れること有り」と読む．江戸中期の後藤芝山は「朋有り，遠方より来たる」と現行のように読む．このように同じ中国の古典文を異なった語順で日本語文語文に読み替えることがある．

(2)「然而」「云爾」の読み方

「然而」をどう読むか．佐藤一斎は「シカクシテ」，後藤芝山は「シカウシテ」，文之玄昌は「カクノゴトクシテ」，明治以降は「シカレドモ」とよく読まれ，明治 25 年以後は「シカリシコウシテ」と読む人が多くなったという．

『論語』述而篇に登場し，和算書の序文やあとがきの締めくくりの言葉である「云爾」を，林羅山や荻生徂徠は現行のとおり「シカイフ」と読むが，その他に「シカリト」「云フコトシカリ」「云フトシカリ」「トシカトス」「云フノミカト」「イフノミ」などの読み方がある．（この項，柳町達也『漢文読解辞典』による）

(3) 返り点をつけない熟語・読まない漢字

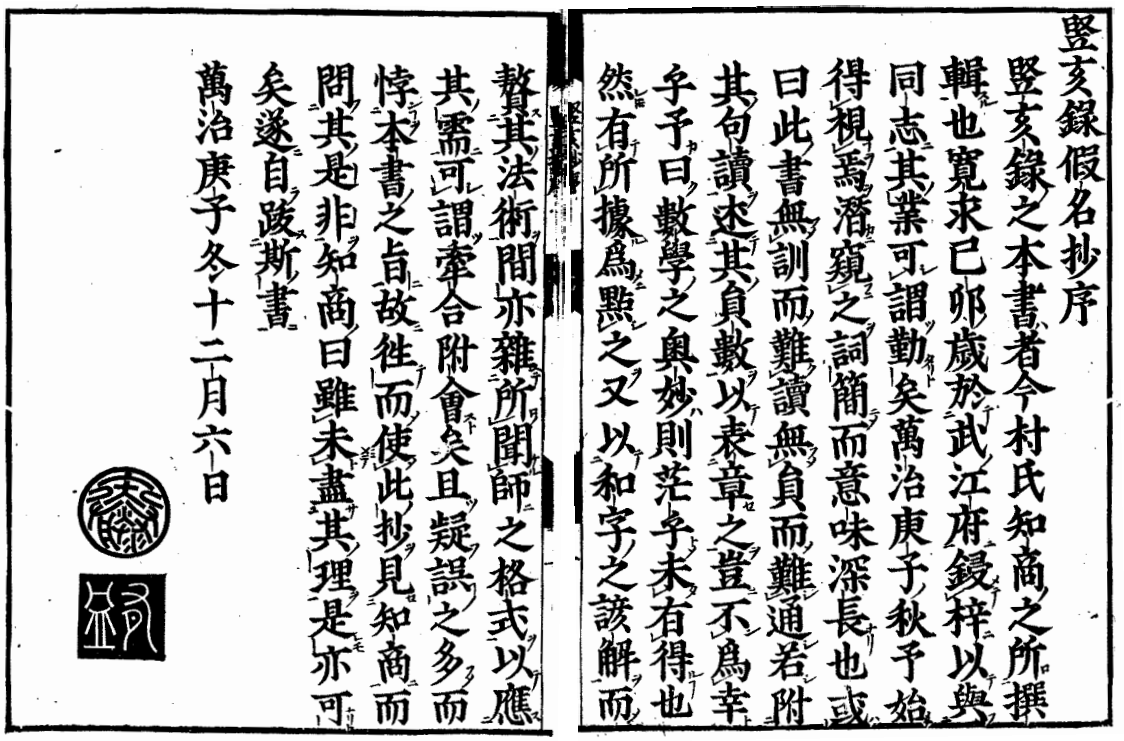
明治 45 年 3 月 29 日の官報記載の「漢文教授に関する調査・訓読法」（服部宇之吉博士）には，「云爾」には返り点をつけず，「しかいふ」と読み，「加之」にも返り点をつけず，「しかのみならず」と読む，と書いてある．

このような異読の例は枚挙に暇（いとま）がない．「則」や「也」を読む人もいれば，まったく読まない流儀もある．『漢文読解辞典』によれば，「寧」は古くから「ムシロ（モシ

ロ）…ンヤ」と読んだが，大正以後「イヅクンゾ」と読むようになった．

（4）竪点の使い方

江戸初期の『竪亥録（じゅがいろく）仮名抄』の序文を示す．『竪亥録仮名抄』は，河内の国の今村知商が漢文で出版した数学公式集『竪亥録』を読みやすいように，弟子の安藤有益が「仮名」に直して解説した本だが，序文は漢文で書いてある．


竪亥錄假名抄序

竪亥錄之本書者今村氏知商之所撰輯也寛永己卯歳於武江府鋟梓以與同志其業可謂勤矣萬治庚子秋予始得視焉潛窺之詞簡而意味深長也或曰此書無訓而難讀無負而難通若附其句讀述其負數以表章之豈不爲幸乎予曰數學之奥妙則茫乎未有得也然有所據爲點之又以和字之該解而贅其法術間亦雜所聞師之格式以應其需可謂牽合附會矣且疑誤之多而悖本書之旨故往而使此抄見知商而問其是非知商曰雖未盡其理是亦可矣遂自跋斯書

萬治庚子冬十二月六日

表題の「假－名」，一行目の「竪－亥－録」，最後の行の「十二－月六－日」など漢字と漢字の間にある短い縦棒のことを「竪点」という．これらの竪点は熟語を示すためのものと考えられるが，なぜか十二月の「十」と「二」の間には竪点がない．

左頁四行目の上のほうには「是非」の二漢字の右側にそれぞれ「竪点」がある．この竪点は「ぜひ」と音読みするよう指示しているようだ．

同じ左頁四行目下の「是＿亦可＿矣（これもまた可なりと）」，四行目の「可レ謂勤＿矣（いいつべし，勤めたりと）」のように漢字の左下に添えた「竪点」は，先行する漢字の「訓読み」を指示しているようだ．是亦可矣の「可」は「よし」と読むのであろう．矣は是亦可矣でも可謂勤矣でも読まない「黙字」になっている．（可レ謂は『論語』雍也篇・子路篇，『土佐日記』や『源氏物語』桐壺に登場し，江戸期では十中八九「いいつべし」と読むが，受験漢文では強調の「つべし」の用法を教えない．）

なお，音読み・訓読みの指示が本によっては左右が逆になっていることがある．

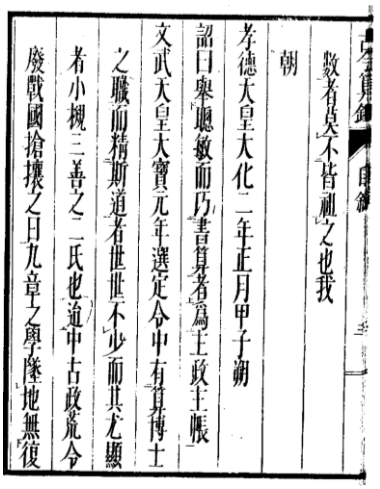
ともあれ竪点は，二漢字の中央にあっては熟語，一つの漢字の左右または左右の下側にあっては音訓の読みを示している．竪点や送り仮名返り点などをまとめて「訓点」と呼ぶ．

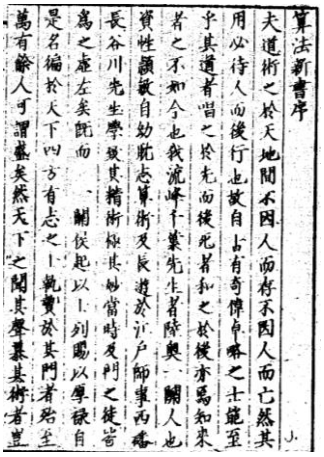
この例文では，右頁七行目（左から三行目）の「表ニ一章之一（これを表章（かきあらわ）せ）」の表章につけられた竪点が「返読には必須の竪点」になっている．なぜなら，「表ニ章

之一」では，「章之表」の語順で読むほかないが，これでは意味が通じないからである．この「表ニ一章之一」のような竪点は，おって述べる「括弧表示」の語順変換では重要な働きをする．ただ非常に残念なことに，『大漢和辞典』や『関孝和全集』など現代に出版された本では，この語順変換に必須の竪点を省略している．

(5) 尊敬語の表記

和算書であれば関孝和，歴史書であれば天皇の名前など，尊敬の念をあらわす最上級の表記法が「台頭」である．通常の文章より一字ないし二字上げて記す．今でいうインデントがマイナスになっている．次級の敬語が「平出」で，古い行は空白のまま残し，新しい行から書き始める．その次のクラスが「闕字（欠字)」で，これは「　　関孝和先生」のように名前の前に一ないし二字分，空白を挿入する．


数者皆吳不皆祖之也我
朝
孝徳天皇大化二年正月甲子朔
詔曰舉聰敏而乃書算者爲上政士帳
文武天皇大寶元年選定令中有算博士
之職而猶斯道者世世不少而其尤顯
者小槻三善之二氏也道中古政亂今
廢戰國捨壞之日九章之類字隆地無復


算法新書序
夫道術之於天地間不因人而存不因人而亡然其
用必待人而後行也故自古有奇傑卓犖之士能至
乎其道者唱之於先而後死者和之於後亦爲知來
者之不知今也我流峰千葉先生者隱奥一關人也
資性頴敏自幼志算術及長遊於江戸師事西磵
長谷川先生學發其精術極其妙當時及門之徒皆
爲之靡左矣既而　關侯起以上列聞以學録自
是名徧於天下四方有志之士競贄於其門者紹至
萬有餘人可謂盛矣然天下之闕真擊暴其術者豈

左の『古今算鑑』内田五観自叙の一行目，「我」から「朝」にかけてが平出，「孝徳天皇」「詔曰」「文武天皇」の三行が台頭している．

右の『算法新書』伊東頼亮序の「既而　　関侯」と二字分空白があるのが，闕字である．

この敬語表記も学校では，まず教えない．

(5) 四声をあらわす記号

学校でも教えないし，漢文研究法や漢文訓読法などの書物にもほとんど書いていない事柄に，それぞれの漢字に記された四声の記号がある．

四声とは，中国語の 4 つのアクセントを示す平声・上声・去声・入声（にっしょう・にゅうせい）を指す．漢詩でいう平仄（ひょうそく）は，上去入をまとめて仄声とよぶことからくる．

林羅山がつけた訓点を道春点というが，道春点の『論語集注（しっちゅう)』には，この四声が説゜のように四隅につけられた小さな丸印で示されている漢字が存在する．右上（説゜）が去声，右下（説。）が入声，左上が上声，左下が平声を示している．

そしてここからが大事なのだが，同じ「説」という漢字でも，「四声によって（アクセントや発音のほかに）意味や品詞が異なる」のである．「説」はふつう入声で読み，「蓋天説」「渾天説」のように名詞であるが，説゜は去声で読み，動詞「よろこぶ」と訓読する．『論語』の巻頭の「学而時習之，不亦説乎」の「また説（よろこば）しからずや」が去声で発音する例になる．

「馬之千里者，一食或尽粟一石，食馬者，不知其能千里而食（馬の千里なる者は，一食に或は粟一石を尽くす．馬を食（やしな）う者は，其の能（のう）の千里なるを知りて食（やしな）はざるなり)」の「一食」の食は入声ショク・シキで名詞「しょくじ」，「食馬」

の食は去声シ・ジで動詞の「やしなう」である.

(6) 訓読における品詞

「東」は，名詞「ひがし」と動詞「ひがしする（東へ向かう)」の二つの訓がある（どちらも平声).「枕」は「まくら」と「まくらする（寝る)」,「欲」は「よく（意欲)」と「ほっする」の二つの訓がある. このような名詞から動詞への変化はしばしば指摘されている. 有名なところでは,「国破山河在，城春草木深」の「春」は動詞で読む.

瓦解（瓦のようにくだける)，蛇行（蛇のようにまがりくねる）のばあいは，名詞が副詞になっている.

形容詞が動詞になる例では,「甘二其食一，美二其服一，安二其居一，楽二其俗一（その食を甘しとし，その服を美とし，その居に安んじ，その俗を楽しむ」. 形容詞の甘と美が「甘いとする」「美しいとする」のように動詞になっている.

ここまでは，学校で学ぶ「品詞の転成」だが，もう一つ，訓読文では「加一三入 寄二乙位一（乙位に寄せるを加入する)」の「寄」のように，はじめは「乙位」を補語とする動詞であり，次の三点に返るときは竪点でつながれた熟語の動詞「加入」の目的語として名詞化している. つまり，二点の「寄」は動詞であり名詞でもある. このように訓読文では英語の ing 形の「動名詞」のような品詞があらわれる.

また，わたしの個人的な意見だが,「未（いまだ…ず)」や「将（まさに…んとす)」などの再読文字は，副詞と中国語文法でいう時間詞をかねた「再読詞」に品詞分類したほうがよいと思う.（「やり残し」の章を見られたい）

4. なぜ訓読がうまれ，発展したか？

言うまでもなく日本の文化は，中国文化を受け入れることからうまれた. 記紀万葉の時代,「やまとことば」を漢字であらわす万葉仮名がうまれたし，奈良平安の時代には仏典を読むため漢字の周囲にヲコト点をつける工夫がうまれた. レ点や一二点などの返り点をつかった訓読もこれらの発展形とみて良いだろう. ヲコト点は各寺院でつけ方が異なったまま伝承され，標準化されたとは言いがたい. 返り点をつかう訓読の表記法は，江戸時代，出版文化の開花とともに，おおむね標準化され，添え仮名・竪点・四声などの細かな工夫がそれぞれの訓読家によって加えられていった.

このような標準化は，武家はもちろん一般庶民への漢文訓読の普及をもたらしたと言えるであろう. 逆に，武家や庶民が生活するうえで欠かせない技能，知識，教養であったからこそ，表記法の標準化が進み，普及が進んだとも言えるであろう.

幕府の歴史書や公文書は漢文で書かれねばならなかったし，各藩の公文書は，候文という変則的な漢文を用いて書かれねばならなかった. 武家階級が漢文を学んだ理由がここにある. 単に儒学を学ぶだけでなく，漢作文のために人々は漢文訓読を学び，研究したのだった.

明治時代に入って，教育的配慮に満ちたこの潮流は明治 45 年の官報告示が「云爾」を「しかいふ」と規定したように，また「有朋自遠方来」を「朋有り，遠方より来る」と読む少数派の読み方が採用されたように，返読の簡素化，訓読文の簡潔化へと進化してゆく.

そしてついに昭和 20 年の敗戦とともに，「漢文」は主要教科の座を追われ，大学入試センター試験に出題されることを唯一の拠り所として，答えが一つしかない「受験漢文」になった．このとき長澤規矩也が重要な役割を果たしたのだが，ここでは省略する．

吉川幸次郎や諸橋轍次など著名な漢文学者の著作を読むと，漢文は意味が通じるように「適当に読めばよい」と書いてある．この自由な訓読法は，返り点の表記が標準化され，南浦文之玄昌や林羅山，佐藤一斎，後藤芝山など諸家が乱立し，それぞれが独自の読み方を工夫した時代の名残りである．あるいは遠く，中国文を日本人が「解釈」さえすればよかった奈良平安の時代の名残りである．「受験漢文」には，吉川幸次郎や諸橋轍次の考え方は通用しない．

5.　わたしの「やり残し」

物理学者のファインマンは，どんなに研究を重ね，成果を得ても，「やり残し」があると述べている．わたしの場合も同様で，和算書の序文を読むことからはじめたが，ほとんどの本に書いていない竪点や闕字・平出・台頭の敬語法，四声のことを知るにつれて，よりいっそう漢文に興味がわいてきた．自然言語は奥が深い．調べれば調べるほど，考えれば考えるほど，新しい発見がある．音楽や絵画などの芸術もそういうものかもしれない．

漢文訓読の標準化された返り点の表記法を数学化する試みは，雑誌「初等数学」に連載してきた．2016 年 1 月末が原稿締め切りの 78 号には，わたしが思いついた漢字 n 字の返り点表記の「場合の数」からなる数列の一般的な母関数の定理と，ダグラス・ロジャース考案の母関数を得る関数方程式族の定理を紹介する予定だ．後者は，有限数学の「組合せ論」と極限概念を結びつける，いわば有限と無限を橋渡しする定理で，世界初公開である．

当時ハワイ大学教授だったダグラス・ロジャースは，オハイオ大学のジェームズ・アンガー教授と一緒にわたしのホームページを読み，2012 年春に「君の発見した数列の母関数は，アトキンソンとシュティットの 2002 年の論文の定理 17 と一致している．初項は君の言うように 1 のほうが良い」と指摘してくれ，おまけに日本大学の斎藤明教授を紹介してくれて，2012 年の応用数学合同研究集会で「漢文訓読の数学モデル」を発表することができた．

アンガー教授は 2013 年秋から国立国語研究所の客員教授になり，11 月には大阪までわたしをたずねてきてくれた．彼は日本語の達人で，奥さんは西宮生れの日本人と聞いた．

ダグラス・ロジャースとは，それからしばらく音信がとだえ，生死のほどもわからなかったが，この 2016 年 1 月に突然「あけましておめでとう．和算のことでちょっと教えてくれ」とメールがきた．当方からは「まだ生きておられるようでご同慶のいたりである」と返事をした．以前もそうだったが，彼のメールは 1 日に 5 通ぐらい矢継ぎ早に届く．

というわけで，雑誌「初等数学」78 号に「ダグラス・ロジャースの極限定理」を紹介する予定のはこびとなった．これが目下の急務となっているわたしの「やり残し」．

この漢文訓読の数学モデルについては，大きな「やり残し」が二つある．

一つは，漢文（中国語の文言文．日本人が中国語のつもりで書いた漢文をふくむ）と訓

読文（日本語の中古文）の，中間によこたわる「返り点つき漢文」の文法構造とくに品詞分類の問題である．「過猶レ不レ及」の元の『論語』の「過猶不及」は中国語，「過ぎたるは猶（なお）及ばざるがごとし」と訓読した文章は明らかに日本語だが，では，返り点のついた「過猶レ不レ及」は何語か？　という問題．中国人に言わせると，「過猶レ不レ及」は変な「レ」の記号が二つくっついた「中国語」になり，「猶」は副詞になるだろうし，日本人は「過ぎたるは猶（なお）及ばざるがごとし」と読んで「日本語」と言い，「猶（なお)」は副詞，「ごとし」は助動詞と言うだろう．

中国語と日本語の中間言語である「過猶レ不レ及」の「猶」は再読詞でいいじゃないか，というのが，わたしの提案だ．

それともう一つ．漢文訓読の数学モデルの応用例を示すこと．数学としての定義や定理は，いまいち世間の人には簡単には受け入れてもらえない．そこで，わかりやすい例を示して，わかりやすい主張をおこなうこと．先従隗始（まず隗より始めよ）．たとえば，江戸時代の漢文がいちばん複雑な構文になったのはいつごろか？　漢文の複雑度のピークは何世紀ごろのことか？　「計量漢文学」とは，こういう疑問を解決する一つの手段である．

6.　計量言語学とは何か

計量言語学では，二つの言語や同一言語の二つの文章が一致・相関するかどうかを，各々から取得したデータ列 $\{x_1, x_2, \cdots x_n\}, \{y_1, y_2 \cdots y_m\}$ をもとにして，統計分析をおこなう．

たとえば，雪を意味する snow（スノー）と Schnee（シュネー），百を意味する hundred（ハンドレッド）と hundert（フンデルト）のように「英語とドイツ語には似通った単語が多い」ことを主張するためには，英独または独英の辞書（ないしは日英・日独の 2 辞書）を用意して，すべての単語または無作為に抽出した一部の単語を調べ，「似通っている」と判定できる単語を書き出してゆく．これらの単語の総単語数に対する比率を計算し，（χ（カイ）二乗検定など）統計学の手法を用いて分析する．この調査は，英語とドイツ語の関係を明らかにするうえで役に立つだろう．

芥川龍之介と谷崎潤一郎の「文章の違い」を検討するには，芥川・谷崎それぞれが書いた文章を大量に集め、たとえば両者の「文章の長さ」（句点（まる「。」）から次の句点（まる「。」）までに含まれる文字数）を調べる方法がある．実際，両者の全集から文章の長さを数え上げ，芥川の文章の長さは谷崎の 2 分の 1 程度，という結果が発表されている．

今日，このような大量の文章を集めたデータベースは「コーパス」と呼ばれている．むろん最近では，インターネットで誰もが参照できるよう，テキストを電子化した「電子コーパス」が主流になっている．

以上のように，計量言語学の調査研究は，一定の大きさをもった語彙表（辞書）やコーパスにもとづいている．語彙表やコーパスから，それぞれの言語や文章を特徴づける（指標となる）データ列を取出し，2 つのデータ列の一致度や相関などの統計分析をおこなうのである．

語彙表やコーパスに含まれる単語には，発音（読み方），品詞分類など，それぞれの単語

のもつ属性が詳しく記されていれば記されているほど，精度の高い分析がおこなえることは言うまでもない．品詞分類では，名詞－固有名詞－地名－日本の地名…のように，第一階層の「名詞」だけでなく，より詳しく分類されているほうが，その文章の特徴をとらえることができる．

これまでの計量言語学の代表的な成果には，安本美典氏による「源氏物語宇治十帖の作者が紫式部かどうか」の研究，大野晋氏による「万葉集から平安時代に至る文学作品に含まれる名詞の比率が直線的に減っている」といういわゆる「大野の法則」の発見，水谷静夫氏による「大野の法則」の修正などがある．安本美典氏は，邪馬台国論争をはじめとする日本古代史や「漢字の将来―漢字の余命はあと 230 年か」などの論考でも知られている．

たとえば，荻生徂徠と新井白石が書いた漢文には「違い」があるのか．江戸初期の林羅山と明治時代の服部宇之吉博士の漢文には「違い」があるのか．どちらがより多く「返読」しているか．使っているレ点の数はどちらが多いのか．頼山陽の『日本外史』と司馬遷の『史記』では，語彙や文体はどう違うのか．何が共通し，何が異なるのか．

7.　考察の対象

日本人の書いた漢文は，「和臭」がするという．それは，日本にしかない国字を使っていたり，中国語の文章とは異なる印象を受けるからであろうが，では誰がいつ書いたどんな漢文が「和臭」の例になるのか，国字の含有率や中国語とは異なる「印象」を具体的に（量的に）示した論考をわたしは知らない．

博士家では読まれなかった助字は，江戸期には訓読されるようになったという．桂庵玄樹に始まり，南浦文之，貝原益軒，太宰春台などにつづく訓読法の変遷には，日本文学史における「大野の法則」のように何か法則があるのか．

論語の「有朋自遠方来」は，「朋有り，遠方よりきたる」と「朋の遠方よりきたる有り」と読む二通りの読み方がある．文之点，一齊点などの相違点を量的にあらわせないか．

和算書のなかには，著者や系統（門流）を明記した直接史料がなく，本文から著者や系統を推定しなければならないものがある．たとえば『求積』は，関孝和の著書かどうか疑われている．むろん，使用されている数学用語や数式の書き方（傍書法），論証・説明の仕方，数値などから，著者や系統が判断できることも多い．間違った記述（たとえば円周率 3.16）は，先行する和算書の影響を受けたものと考えられる．

和算書の漢文に限らず，日本人が書いた漢文には，送り仮名，読み仮名（右訓・左訓），圏点・傍点（批点）の句読点，音と訓の区別や熟語を示す竪点（たててん．立点，縦点とも．年賀パズルの景品としてこの「計量漢文学への道」を書く前は「熟字訓と竪点の研究」を書くつもりでいた．熟字訓は「仮如」を「たとえば」，「已矣哉」を「やんぬるかな」と読むようなこと）などが添えられることもある．また，闕字・平出・台頭などの敬語表記法が用いられることもある．

ここでは，これらの添え字（「或問」を「或る人問う」と読むとき，漢字「人」を添える．繰返し記号なども添え字の一種），句読点，敬語表記法を無視し，返り点の表記と返り点が

もたらす語順変換（返読）だけを考察の対象とする．

そもそも何故，中国文に返り点をつけて日本語として読む，という伝統が生まれたのだろうか．もしかすると，江戸時代には「返り点をつけることができるように訓読するようになった」のかもしれない．

ここでは，「返り点つき漢文」を「括弧表示」することによって，統計分析の対象となるデータ列を取り出すいくつかの手法を紹介する．

8. 括弧表示とは何か

括弧表示では，α＜β＞のように，漢文を返読するとき，先に読むβの部分（β項）を括弧＜　＞でくくり，後から読むα（α項）を左括弧＜の左側に置く．αは漢字 1 字の場合もあれば，竪点（たててん＝ハイフン）－で結ばれた 2 字以上の漢字の場合もある．返読は入れ子になることもあるので，β 項のなかにα′＜β′＞のような部分があってもよい．

代表的な例を下に示す．中央が括弧表示，右側は返読した結果の漢字列をあらわす．

我心匪レ石不レ可レ転也 → 我心匪＜石＞不＜可＜転＞＞也 → 我心石匪転可不也

師不二必賢一於弟子一 → 師不＜必賢＜於弟子＞＞ → 師必於弟子賢不

不下以二千里一称上也 → 不＜以＜千里＞称＞也 → 千里以称不也

如欲三平二一治天下一 → 如欲＜平一治＜天下＞＞ → 如天下平治欲

此非四吾所三一以居二一処子一也 → 此非＜吾所一以＜居一処＜子＞＞＞也

→ 此我子居処所以非也

括弧表示の利点は，A一BC二，A上BC下，AレB－C（3 つとも BCA の順に読む）のような異点同順の表記を統一し，しかも＜，＞，－の 3 つの全角サイズの記号を漢字列につけ加えるだけで「返り点が指示する語順変換」を表現できることにある．むろん，縦書・横書のどちらにも対応している．再読文字（□で囲む）は，

過猶レ不レ及 → 過猶－＜不＜及＞ → 過猶及不猶

のように「－＜」で扱える．（ごく稀に「再読文字の連用」があるが，ここでは省略する）

再読をふくめた数学モデルは，2012 年の計量国語学会で発表したが，以下の議論では，再読は無視し，品詞分類をおこなうときに再読文字を「再読詞」と分類することにする．

9. 返読・再読パターンの頻度

漢字 1 字には，返り点のつけようがない．訓読パターンは漢字 1 字をそのまま読む 1 通りしかない．

漢字 2 字のばあい，少年（AB）有レ朋（A＜B＞）のようにレ点の有無により，2 通りの訓読のパターンがある．

漢字 3 字では，我独醒（ABC）可レ妻也（A＜B＞C）朝聞レ道（AB＜C＞）不レ踰レ矩（A＜B＜C＞＞）楽二其俗一（A＜BC＞）の 5 通りのほか，竪点を使った卑二一下之一（A－B＜C＞）を加えた 6 通りがある．

このようにn 字の漢字列に対する，＜ と ＞の対および－のつけ方の総数k_nは，$k_1 = 1, k_2 = 2, k_3 = 6, k_4 = 20, k_5 = 70, k_6 = 254, \cdots$となる．（竪点をつかわない場合，訓読パターンの数列はカタラン数になる．カタラン数の初項c_0は 1 にする．数列$\{k_n\}$の初項k_0もカタラン数の拡張という意味で 1 にしたほうがよい．前出のダグラス・ロジャースの最初のメールを参照）

入れ子になる＜　＞のうち，もっとも外側にある＜　＞を考えると，n 字の漢字列$A_1A_2 \cdots A_n$に対する返読・再読のパターンは，A_1＜$[k_{n-1}]$＞のようにk_{n-1}通りあるもの，$A_1 - A_2$＜$[k_{n-2}]$＞のようにk_{n-2}通りあるもの，…，$A_1 - A_2 \cdots - A_{n-1}$＜$A_n$＞のように$k_1 = 1$通りあるものがある（[　]はβ 項の場合の数を示している）．すなわち n 字の漢字列に対する，「最長」の返読・再読のパターンの総数は$\sum_{i=1}^{n-1} k_i$になる．よって，十分に大きな n までの n=2,3,…のパターンの合併集合が「元の返り点つき漢文で使われている返読・再読パターン」といえる．一つの「返り点つき漢文」のなかで，これらのパターンが各々何回使われているかを数えあげれば，データ列$\{x_1, x_2, \cdots x_k\}$が得られる．

10.　bigram によるデータ列の取得

漢字と記号を同時に扱う手法のひとつに，bigram（バイグラム）がある．一般に，漢字と記号あわせて N 文字の列の最初の 1 字から順に 1 字ずつずらしながら n 文字を取り出す手法を n-gram とよぶ．（全部で$N - (n - 1)$個のデータが取り出せる）

括弧表示の bigram には，

漢漢，漢－，漢＜，漢＞，－漢，＜漢，＞漢，＞＞

の 8 パターンがある．これらの総数をカウントし，データ列$\{x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8\}$とする．

ただし，この数え方では，たとえば①ABC＜D＞と②A＜BCD＞では（A, B, C, D は漢字），

①　漢漢 2 個（ＡＢ，ＢＣ），漢＜1 個（Ｃ＜），漢＞1 個（Ｄ＞），＜漢 1 個（＜Ｄ）

②　漢漢 2 個（ＢＣ，ＣＤ），漢＜1 個（Ａ＜），漢＞1 個（Ｄ＞），＜漢 1 個（＜Ｂ）

のように，①②の個数がまったく同じになる．じつはこのようなケースは，バイグラムでは常に発生することが証明できる．

「漢字 2 字」とその前後・中間の「＜，＞，－」記号をふくめてパターンを数える手もある．これは，たとえば，我心匪レ石不レ可レ転也の「我心」のように漢字２字から始まる部分であれば，漢漢，漢漢＜，漢漢＞，漢漢＞＞，漢漢＞＞＞…を区別して数えるが，これとて厳密に言えば，上の①②のようなことが避けられない．

なお，このように分類したパターンに表われる漢字の「品詞分類」を加味すれば，より大きなデータ列が得られる．

以上紹介した 2 つのデータ取得法は計画にとどまっており，今後，どのような統計分析の手法（分類器）を用いれば，どの程度の精度が確保できるかを知るには，理論的考察とあわせて実験を重ねるほかない．（「計量漢文学への道」終り）