

## パスカル全データベース作成と言語解析 (III)

白石修二<sup>†</sup> 輪田裕<sup>†</sup> 藤村丞<sup>†</sup>

福岡大学 理学部 応用数学科<sup>†</sup>  
福岡大学 人文学部 フランス語学科<sup>†</sup>

メナール版パスカル全集を底本として、第2巻のパスカル頻度表 2-1 を作成した。2-1 は、第2巻でパスカルが書いた部分を抜粋した。第2巻は収録作ごとに、アルファベット順、頻度順、逆引き順、と作成し、次に、全収録作を併せての全体で、アルファベット順、頻度順、逆引き順を作成した。第2巻のパスカルが書いた部分のテキストデータベースは、この頻度表の完成により、その精緻性に関してかなり完成度が上がったことを意味している。

また、同時に平行して、第3巻、第4巻のすでに電子化したテキストをより信頼性の高いデータベースとするために、第2巻と同様に収録作ごとの単語の頻度表を作成しながら、テキストデータベースを検証中である。

## Construction of the Database on the Complete Works of Blaise Pascal and Linguistic Analysis (III)

SHIRAISHI Shuji<sup>†</sup> WADA Yutaka<sup>†</sup> FUJIMURA Shou<sup>†</sup>

*Department of Applied Mathematics, Faculty of Science, Fukuoka University<sup>†</sup>*  
*Department of French Language, Faculty of Humanities, Fukuoka University<sup>†</sup>*

We wrote up a word frequency book 2-1 on the second volume of the complete works of Blaise Pascal on the Jean Mesnard version as definitive version. This word frequency book 2-1 is composed by the tables of every work of Pascal in alphabetical, frequent and reverse orders, and by the tables of whole work of Pascal in same orders. We could construct a more complete text database of the second volume by the completion of this word frequency book. We continue constructing the database of the third and fourth volumes through examining the word frequency to improve the completeness of text database.

## 1 はじめに

昨年度来、引き続きパスカルのデータベース作成を行っている。メナール版パスカル全集 [1] を底本とし、既刊されている全 4 冊、第 1 巻から第 4 巻までをすでに電子テキスト化し、頻度表作成により、信頼性の高い、より精緻なデータベースを構築中である。

第 2 巻のパスカル自身が書いた部分を抜粋しての電子テキストは、パスカル頻度表 2-1 の完成により、データベースとしての完成をみている。2-1 は、第 2 巻でパスカルが書いた部分のみを抜粋した。頻度表は、まず 2 巻中の収録作ごとに分けてアルファベット順、頻度順、逆引き順と作成し、また同時に、メナール版の特徴として、時系列的に作品を並べていることから、全体を併せて一つの作品と考え、この頻度表も作成した。これについても同じようにアルファベット順、頻度順、逆引き順と作成した。

この頻度表の作成は、誤入力の見逃しに貢献し、データベースをより完璧なものに近づけるものである。と同時に希少語の見逃しなどにみられるように、1 次的な資料としての役割も担うので、大いに意義のあることとして、細心の注意をはらって取り組んできたものである。

このパスカルデータベースは、メナール版に忠実に電子テキスト化している。このことは、パスカル本人が書いた部分と、それ以外の部分（例えば引用、パスカル宛ての書簡、メナール自身の脚注）などが混在しているメナール版では、通常の方法では、その頻度の現れ方に、パスカル自身の言葉ではない言葉も混入してくるのは避けられない。そこで、パスカルの言葉のみを自動的に切り出せるように目印を入れ、パスカル自身の言葉にこだわり、細心の配慮をもって作成した。メナール版にいくつかのスペルミスがあったことを附記しておく。

## 2 データベース 2-1 の完成

第 2 巻には、1640-1654 年代の 34 の作品が収録されている。パスカル初期の作品で、文学や哲学に留まらない数学や物理の科学論文を含む。データベース 2-1 は、第 2 巻でパスカルが書いた部分を抜粋したものであるが、作者に関して明記してはいないが、パスカルの作品と認められる可能性の高いものは 2-1 に含めた。結局、パスカルの作品として収録したデータベース 2-1 には 32 の作品を収録した。

データベース 2-1 の作品内容と大きさ (バイト) は以下である。

**2-228-235** 『円錐曲線試論』 1640. 初. 最初の幾何学論文。後にこれを完成させるべく努める。(7498byte)

**2-282-283** 姉に宛てた手紙。1643.1.31. 現存する (2782byte) 最初の手紙。

**2-311-312** パスカルのものとされる詩。当時の記録作家 Tallemant des Réaux によって書き留められた。メナールは、詩人としての妹に対する対抗意識の可能性を指摘している。(1184byte)

**2-331-341** セギエ大法官への献辞とその使用方法。1645. パスカルは計算機を完成させるが、セギエはそのさいに力添えをしてくれた。(25097byte)

- 2-498-508 『真空に関する新実験』1646.10.8. ルアンで行ったトリチェリの実験追試と彼の工夫による他の実験にかんする報告。パリに戻ってから出版。(25183byte)
- 2-518-527 ノエル神父に宛てた論争の手紙。1646.10.29. 小冊子『真空に関する新実験』に対し、ノエルは彼に真空の存在を否定する手紙を書く。これは、それに対する反論。真空論争と呼ばれる。(22344byte)
- 2-553-555 姉に宛てた手紙。1648.1.26. パリのポール・ロワイヤル修道院での出来事を伝えている。(6360byte)
- 2-559-576 数学愛好家であるル・パユールに宛てた手紙。1648.2. ノエル神父の再反論に直接彼は反論しなかったが、内容は真空論争の延長。(43829byte)
- 2-581-583 彼と妹が姉に宛てた手紙。1648.4.1. 筆跡は妹のものだが、口述筆記の可能性もある。きわめて宗教的内容。(7589byte)
- 2-677-690 『流体の平衡についての大実験談』1648. 秋. パスカルが真空の実験を生地クレルモンにあるピュイドームの麓と山頂で行った気圧の実験報告。実験そのものは義理の兄が行い、彼からの報告をそのまま挿入しているが、この表では、それを省いている。(29413byte)
- 2-695-698 彼と妹が姉に宛てて書いた手紙。1648.11.5. 筆跡は妹。ただし、最後の一行だけがパスカルの筆跡。はたして、これが口述筆記なのかどうか不明。きわめて宗教的内容。(9440byte)
- 2-705-706 パスカルから姉に宛てた手紙。1649.3. 終。(1082byte)
- 2-712-715 パスカルの計算機への国王の特許状。1649.5.22. メナールによれば、この一部分はパスカルが書いたものを正確に引き写した。ここでは、その部分に該当すると思われる個所だけを対象とした。(7757byte)
- 2-777-785 『真空論』序文断片。1651. つづく『真空論』断片とともに、『真空に関する新実験』で予告された『真空論』のための断片。(16200byte)
- 2-787-798 『真空論』断片。1651.(21673byte)
- 2-804-813 リベイル宛ての手紙。1651.1.12. クレルモンフェランにあるジェズイットの学院に提出された哲学の学位論文が、真空実験におけるパスカルの優先権を否定したことに対し、パスカルは異義を唱えた。(22979byte)
- 2-817-818 リベイルの返事に対するパスカルの手紙。1651.8.(4527byte)
- 2-843-844 父の墓碑名。1651. 終. 父エチエンヌは1651年9月24日に亡くなる。この墓碑名はパルカルの作と思われるが、確証はない。(1726byte)
- 2-851-863 父の死について、義理の兄に宛てて書かれた手紙。1651.10.17. (27302byte)
- 2-923-926 計算機をスエーデンの女王クリスティーヌに献呈する手紙。1652.6. (7008byte)
- 2-934-935 義理の兄がクレルモン近くに館を購入したさいの契約書。ここにパスカルは冬を過ごしている。1652.11.6.(1142byte)
- 2-941-944 パスカルの御者が地中に貯めていた財産を、パスカルから手渡すための証書。1653.3.5.(2784byte)

- 2-951 パスカルから義理の兄に宛てた手紙。1653.6.6.(1087byte)
- 2-1031-1035 『いとも高名なるパリ数学アカデミーに』1654.パリのアカデミーに提出された研究計画書。ラテン語で書かれている。(5948byte)
- 2-1041-1101 『流体の平衡と大気の重さについて』1654。「流体の平衡論」「空気塊の重さ論」「前二論文の結論」の三部からなる。真空論の発展と完成。死後出版。(142587byte)
- 2-1108-1119 『円錐曲線の発生』ライプニッツが書き写したものであるが、ほぼ、正確に写しているものと思われる。パスカルが16歳のときに書いた『試論』の発展であるが、その第一章のみが写本として残っている。原文はラテン語。(11047byte)
- 2-1137-1144 パスカルからフェルマーに宛た手紙。1654.7.29. 数学、特に賭の分け前に関する手紙。部分的に、特に数学的論証部分にラテン語が使用されている。(16717byte)
- 2-1147-1153 パスカルからフェルマーに宛た手紙。1654.8.24.(15431byte)
- 2-1158 パスカルからフェルマーに宛た手紙。1654.10.27.(1054byte)
- 2-1176-1287 『算術三角形論および付属論文』1654. ラテン語で書かれている。(105958byte)
- 2-1288-1329 『算術三角形論』『単位数を母数とする算術三角形の様々な応用』『数序列論』1654. フランス語版。(103524byte)
- 2-1329-1332 『数序列論』1654. フランス語版に付随しているが、ラテン語で書かれている。(3275byte)

### 3 パスカル頻度表 2-1 の完成

パスカル頻度表 2-1 の主な内容は以下である。

1. 作品ごとの単語の頻度データ (アルファベット順, 頻度順, 逆引き順)。
2. 全作品中の単語の頻度データ (アルファベット順, 頻度順, 逆引き順)。
3. ラテン語作品の頻度データ。

### 4 パスカル頻度表作成上の留意点

各巻の頻度表の作成基準として、作品ごとに万全を期した頻度表を作成し、それをもっと大きく基準を定め、一つの定義によって集成し、最終的にはパスカルが書いた部分とそれ以外の部分との2つに大別し調べることを目的とする。例えば、第2巻の最終目的に関しては、パスカルが書いた部分を2-1、それ以外の部分を2-2とまとめることにある。

頻度順の章では、その言葉が全体の語彙に占める割合を示した相対頻度も付加してある。完全なラテン語の作品に関しては、フランス語と分離して扱った。物理や数学の作

品での数学記号や数式は1つの単語として扱っている。省略形には、原形を補っている。倒置形として使われている連続した単語は、1つの単語とした。なるべく1次資料的な頻度表作成に徹した。

## 5 おわりに

1997年度重点領域「人文コンピュータ」と1998年度特定領域「人文コンピュータ」に連続して採用されたことによって、メナール版パスカル全集を底本とし、第1巻から第4巻までのデータベース化が強力に推進され、今なお現在形で進行中である。

今回の成果としては、第2巻のパスカルが書いた部分のデータベース 2-1 と 頻度表 2-1 の完成をみる事ができたことである。

また、文部省の学術図書の助成により、パスカル頻度表 2-1 を 1999年3月に刊行する運びとなり、多くの研究者にこの貴重な資料が供されることとなったのは嬉しい。

一連のデータベース作成上の作業に関しての手順、留意点、など全ての仕事に関しては、1998年度の報告集で詳述する予定である。

これら一連の作業が終了次第、メナール版がまだまだ完結をみないことを考えるとオリジナル原稿のデータベース化を試みることも必要かと考えている。

## 6 パスカル頻度表 2-1 の サンプル

2-498-508 の作品からの、一部抜粋である。

### 6.1 アルファベット順

expérience	4	grain	2	honneur	1
expériences	14	grand	21	horizon	3
extraordinairement	1	grande	5	horreur	8
extérieur	2	grandeur	1	hors	2
< F >		< H >		huile	1
facilement	1	hasard	1	< I >	
faiblesse	1	hausse	1	ici	1
faire	17	haut	10	il	45
faisant	3	haute	2	< J >	
faisons	1	hauteur	31	j'	16
fuyait	1	hermétiquement	2	jambe	9
< G >		heureuses	1	jambes	4
gens	1	hommes	1	je	18
Gesvres	1	honnêtes	1	joignent	1

jointe	3	justes	1	laissé	1
journalières	1	< L >		laiton	1
journellement	1	l'	121	laquelle	11
jugé	1	la	118	le	125
jusques	14	laissant	1	lecteur	2
juste	2	laisse	5		

## 6.2 頻度順

178	de	29	pour	13	comme
144	et	29	se	13	corps
125	le	28	ce	13	ont
121	l'	27	pas	13	plein
118	la	25	cette	13	sorte
106	que	25	une	13	tous
81	en	24	par	12	apparent
72	à	21	deux	12	ces
71	les	21	espace	12	dont
70	qui	21	grand	12	ouverture
70	qu'	21	tuyau	12	peu
70	un	19	même	11	a
64	d'	18	je	11	air
61	dans	17	autre	11	au
61	vide	17	faire	11	bas
54	est	17	ne	11	c'
49	on	16	elle	11	celle
45	du	16	j'	11	laquelle
45	il	16	nature	11	où
44	eau	16	ou	11	tout
40	n'	15	apparence	11	toutes
37	si	15	vaisseau	11	verre
37	vif-argent	14	ai	10	cet
34	plus	14	étant	10	doigt
33	avec	14	expériences	10	environ
33	y	14	force	10	haut
32	des	14	jusques	10	seringue
31	hauteur	14	piston	10	soit
29	pieds	14	sont	9	admettre

### 6.3 逆引き順

a	donc	ce
la	etc	surface
cela	pied	place
confirma	demi-pied	espace
voudra	grand	répugnance
sera	quand	France
divisera	descend	connaissance
arrivera	rend	puissance
trouvera	comprend	impuissance
pourra	tend	substance
sa	attend	avance
divulgua	répond	croyance
éprouva	hasard	expérience
avec	jambe	violence
public	tombe	

### 参考文献

- [1] Blaise Pascal, Œuvres complète, DDB.
- [2] 輪田, 白石, 柴田, パスカル頻度表 2-1 (多賀出版, 1999年3月出版予定)
- [3] 有川ほか, テキストデータベース管理システム SIGMA 第2版について, 九大大型計算機センター広報, Vol.20, No.6 (1987)
- [4] パスカル全データベース作成と言語解析 (I) (白石, 輪田, 藤村), 情報研報, Vol.97, No81, 31-36 (1997)
- [5] パスカル全データベース作成と言語解析 (II) (輪田, 白石, 藤村), シンポジウム, 人文科学における数量的分析 (III) (1998)
- [6] 1997年度重点領域(人文コンピュータ)研究報告集