報告区分 1

テキストマイニングを利用した群棟型集合住宅における設計者言説の特徴分析

Feature Analysis of Architect's Discourse in Group Housing by Text Mining

○李尚勲[*1]、服部岑生[*2]
LEE Sanghoon, HATTORI Mineki

This is a study on the feature of the architect's discourse about “Group Housing ” by the method of text mining. “Group Housing ” is composed of two, three or sometimes more buildings.
We focus some typical and specific words that appear frequently and partially depending on the housing group, and analyze the relationship between the number size of building and the usage of these words.
As a conclusion, the housing which is more than two buildings is described using more frequently by the words responding to the context of site. It defines that “Group Housing ” has potentiality for the improvement of urban environment in the case of respective residential renewal.

キーワード：都市型集合住宅，群棟型，言説，テキストマイニング，個別更新
Keywords: Urban Housing, Group Housing Type, Discourse, Text Mining, Respective Renewal

１．研究の背景・目的

東京の今後の住宅建設は、住宅市街地の拡大でなく、既成市街地での更新によると言われている。また、日本の都市はヨーロッパとは違って、建築行為の自由度が大きく、所謂「巨大な集落」や「私有地の集積体」とも言われている[文1)]。老朽化した既存住宅ストックの有効活用や個別更新のための住宅計画の確立は、都市住居としての性能確保のみならず、望ましい住宅地形成への大きな役割を担っている点から、都市住居計画の主要な研究分野の一つであるといえよう。

一方、服部ら[文2)]が言うように、中小規模の敷地で住宅の更新を進めるとすると、普遍的な経済活動の観点からは、容積や建設権利の上限まで、ますます高層化・高密度化していく住宅計画が予想される。

しかし、近年の集合住宅ブームと相まって、日本における都市型集合住宅の計画傾向を俯瞰すると、経済性重視や住戸内部の性能確保目当てのみからは説明が取れない住棟容積の低密度さや群棟によるデザイン行為といった計画傾向がよく見られる。

筆者は、前稿[文3)]で現代日本における小規模都市型集合住宅に見られる複数住棟による群の構成＝群棟性に着目した。複数の住棟に現れた計画の共通の特徴に注目し、群棟の様子について考察した[注1)]。

本論では、前稿[文3)]の住棟数の分類に従い単棟型と「群棟型」（2 棟以上の事例で、前稿[文3)]の分棟型）とし分析した。住棟の群の構成には、設計者の示し合わせはないが、共有する価値観や計画意図があると仮定する。

また、それに関する情報は設計者自身の解説文に直接的・間接的に表れると考え、表明された建築データ諸元を利用するが建築作品そのものの計測とその計量分析でなく、テキストマイニング[注2)]でその設計者言説を解析し、住棟数との関連で表れる設計者言説の共通的な特徴を明らかにすることを目的とする。この成果は、個別更新により変化していく住宅市街地において、群棟型集合住宅が持つ意義と可能性を推理するのに有効な知見を与えると考えた。

*1　NPO 法人　ちば地域再生リサーチ、研究員、工学修士　Researcher, Chiba Regional Revitalization Research, M.Eng.
*2　千葉大学、名誉教授、工博　Prof. Emeritus, Chiba University, Dr.Eng.

２．本論の位置付け

近年計画されている都市型集合住宅に対する設計論、とくに群棟と類似概念としての小規模集合住宅群や周辺への影響・個別更新のあり方などについては、服部ら[文2)]の研究がある。また、建築家の言説を研究したものとしては、奥山ら[文4)文5)文6)注3)]の住宅を対象とした一連の研究がある。テキストマイニングを用いた研究としては、テキストマイニングでアンケート調査結果を分析した研究[文7)]、建築評論書を分析した研究[文8)]があるが、新種の建築計画の特徴を見いだしていくような本稿の研究は、新しい試みである。また、本論は、設計者の言説を読みとる定量的方法として優れた研究方法の一つであるテキストマイニングを用い、小規模都市型集合住宅、とくに群棟型集合住宅計画に関わる設計者言説の特徴を明らかにする最初の研究である。

３．研究方法・概要・対象

３－１．研究方法・概要

本論では、二つの研究仮説を検証する。一つ目は、計量言語学分野では日本語で書かれたテキストの場合、高頻度語[注4)]や広範囲語[注5)]は常用性の高い語群である可能性が高いと言われているが、小規模都市型集合住宅の設計者の言説から抽出した高頻度語と広範囲語もその仮説通りであるかどうかを基本分析で検証する。

二つ目は、群棟型集合住宅の設計者言説は、群棟の住棟数が違うにつれて言説の内容も違うという仮説を立て、住棟数別グループ毎から特徴語を抽出する特徴分析で検証し、また、対応分析により各グループと特徴語との関係を視覚化し確認する。但し、元の言説全文をすべて調査するのは非効率的であるため、言説の内容を代表する概念を客観的に抽出する必要性がある。そこで、本論では元の言説を形態素解析[注6)]を通じて分割し、各単語の品詞情報の中、名詞のみを抽出して分析を行う。

研究の流れとしては、①収集した言説の自然言語処理[注7)]を行い、得られたテキストデータに対して形態素解析を通じ、品詞の名詞であるキーワードのみを抽出する。

②基本分析：高頻度語や広範囲語の出現傾向から、一つ目の仮説を検証する。抽出キーワードの全般的な構造も理解する。

図 1 研究の流れ

③二つ目の仮説の検証(特徴分析)：住棟数別グループ毎の特徴的なキーワードを抽出分析する。

④対応分析：各住棟数別グループと特徴語との関係を視覚化し理解を深める。

⑤分析結果を総括し、成果と課題について言及する。

３－２．研究の対象

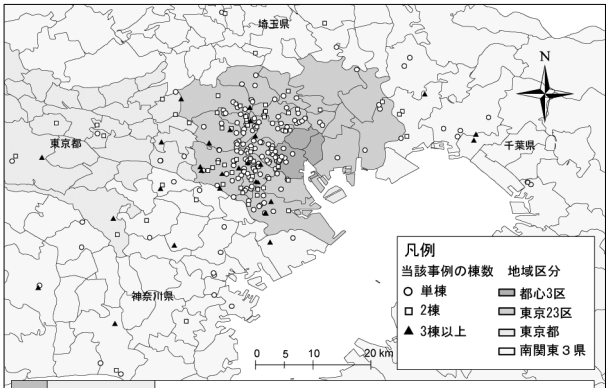
（１）研究対象：事例

1980 年から 2007 年まで、建築専門誌に紹介された集合住宅の中、南関東市街地を中心とした小規模である事例を対象とする。住棟数による群棟デザインの分類は、前稿に倣う[注8)]。また、グルーピングは、単棟と 2 棟、そして 3 棟以上の三つのグループに分けて、該当事例の言説を分析する[注9)]。

（２）研究対象：言説

収集した総言説数は 260 個(事例数 272)である[注10)]。

さらに、一つの言説に異なる住棟数の事例が混じっている 4 個の言説を除き、結果、256 個の言説（262 事例）を分析対象とする。



調査範囲	建築ジャーナル	新建築(1980年1月号～2007年10月号) 住宅特集(1985年2月号～2007年10月号)		
	地域	南関東1道3県		
	建築規模	敷地面積	階数	建築規模
		1,000㎡未満	5階建て以下	共同住宅、長屋、寄宿舎・寮

群棟数グループ別言説数(総言説数)	単棟	2棟	3 棟 以 上				
256(260)			3棟	4棟	5棟	6棟	7棟
群棟数別言説数(事例数)	167(172)	62(62)	20(21)	3(3)	2(2)	1(1)	1(1)
平均敷地面積(㎡)	335.63	456.10	642.65				
平均建蔽率&容積率(%)	58.09・162.87	55.18・134.27	51.59・126.20				
平均許容建蔽率&容積率(%)	65.45・190.18	59.40・163.38	57.80・155.67				
平均住戸数(戸)	8.33	11.13	14.68				
群棟分類例 左から：単棟(ベルム経堂)、2棟(コルテ松波)、3棟以上(エクセル米喜)	単棟(ベルム経堂)	2棟(コルテ松波)	3棟(エクセル米喜)				

図 2 研究対象(事例・言説)の概要[注11)]

４．前処理

４－１．分析ツールおよび前処理の仕方・経緯

分析は、MeCab[注12)] (ver. 0. 994)を利用して言説のテキストデータの形態素解析を行い、その中から品詞情報を

基準とし、キーワードを抽出する[注13)]。

また、言説が建築専門ジャーナルに掲載した記事であることを考慮し、使用者辞書を作成し登録する[注14)]。事前にクリーニング作業も済ませておく[注15)]。

４－２．調査品詞の限定と抽出

形態素解析を行った結果、253,175 語が抽出されたが、基本辞書で使われている品詞体系の中から、名詞のみを分析対象として選ぶ[注16)]。

表 1 は事例として「Nakamaru House」の言説とそれを形態素解析し、名詞のみを抽出した結果の一部を示す。

表 1 形態素解析結果の例

事例 no1. Nakamaru House(掲載誌：住宅特集 85win /設計：鈴木エドワード)

波(wave)の実態。 波は自然におけるひとつの根本的な力である。波は時には肉眼で見ることができるが、その実態をつかむことはできない。波は時に乗って空間を駆ける。まるで生命のように…。人間は光なしには何も見ることができないが、その光は微粒子(particles)、波(waves)、またはそれらの複合体(wavicles)で構成されているという。	波 I wave I 実態 I 波 I 自然 I 根本 I 力 I 波 I 肉眼 I 実態 I 波 I 空間 I 生命 I 人間 I 光 I 光 I 微粒子 I particles I 波 I waves I 複合体 I wavicles I 構成

５．基本分析

５－１．分析結果の概要

総延べ語数[注17)]は元の形態素解析情報量の約 27%である、67,799 語が抽出された。総異なり語数[注18)]は 8,968 語が抽出された。キーワード[注19)]の頻度分布を見ると、頻度 100 未満のキーワードが抽出結果の約 99%を占めている。とくに、頻度 1 のみのキーワードが全体の約 47%を占めており、キーワードの半分近くは一つの事例でしか使われていないことが分かる[注20)]（表 2)。

表 2 基本分析結果の概要

階級(頻度)	キーワード数	構成比
～100	8,872	98.93%
～200	65	0.72%
～300	15	0.17%
～400	7	0.08%
～500	2	0.02%
～600	3	0.03%
～700	2	0.02%
～800	1	0.01%
～900	0	0.00%
～1000	0	0.00%
～1100	1	0.01%
合計	8,968	100.00%

	全グループ
延べ語数	67,799
異なり語数	8,968

全テキストデータ概要(名詞)

	グループ		
	単棟	2棟	3棟以上
延べ語数	43,064	17,680	7,055
異なり語数	6,946	4,172	2,390

住棟数グループ毎のテキストデータ概要(名詞)

階級(頻度)	キーワード数	構成比
1	4,178	46.59%
～10	3,563	39.73%
～50	986	10.99%
～100	145	1.62%
～1100	96	1.07%

５－２．高頻度語・広範囲語の出現傾向

表 3 は高頻度語と広範囲語の上位 100 位までのリストの一部である。全般的に両方とも上位に上るほど、同じキーワードが置かれている。元のテキストデータの抜粋先が建築専門誌であることを勘案すれば､「空間」「建築」「建物」「計画」（左から高頻度 1・2・5・9 位)等、建築計画と密接した関連用語が上位にランクしていることから、調査対象のテキストの高頻度語と広範囲語も無特徴語的な性格が強いと言えるであろう。

また、「住戸」「住宅」「集合住宅」（左から高頻度 3・4・8 位)などのキーワードが上位にランクしているのも、集合住宅を計画した設計者の言説であることを思案すればごく自然な結果であると思われる。

「都市」「敷地」「道路」「街」（左から高頻度 6・7・16・26 位)などが上位にランクしていることから、集合住宅計画を取り巻く環境に対して、多様なレベルが語られていることが推察できよう[注21)]。

表 3 上位 100 位までの高頻度語・広範囲語の一部

高頻度語上位リスト		
Rank	キーワード	頻度
1	空間	1,086
2	建築	724
3	住戸	672
4	住宅	667
5	建物	568
6	都市	548
7	敷地	540
8	集合住宅	495
9	計画	418
10	壁	376
11	構成	361
12	部分	329
13	環境	323
	生活	323
15	設計	319
16	道路	300
17	構造	299
18	賃貸	298
19	家	283
20	m	273
21	住居	258
22	可能	253
23	人	238
24	階段	235
25	関係	233
26	街	231
27	場所	227
28	中庭	222
29	開口	204
30	ユニット	203
[illegible]	[illegible]	[illegible]
89	[illegible]	105
90	居住	103
	状況	103
92	位置	102
93	かたち	101
	バルコニー	101
95	メゾネット	100
	共用	100
	廊下	100
98	アプローチ	99
	表現	99
100	土地	98

広範囲語上位リスト		
Rank	キーワード	出現言説数
1	敷地	201
2	空間	200
3	建物	180
4	住宅	169
5	建築	152
6	集合住宅	150
	住戸	150
8	計画	149
	構成	149
10	生活	137
	壁	137
12	道路	132
	部分	132
14	賃貸	130
15	可能	119
	環境	119
	設計	119
18	階段	117
19	構造	116
	都市	116
21	m	113
	配置	113
23	周辺	109
24	内部	104
25	住居	103
26	外部	101
27	開口	100
28	[illegible]	99
[illegible]	[illegible]	[illegible]
87	家	58
88	住まい	57
89	バルコニー	56
	家族	56
	表現	56
	要素	56
93	アパート	55
	ボリューム	55
	居住	55
	形態	55
	十分	55
	鉄筋コンクリート	55
	風景	55
100	共有	54

一方、「中庭」の場合、高頻度語の順位は 28 位(頻度：222)であることから、小規模都市型集合住宅の計画に際して、設計者の言説から多用される用語の一つであると読みとりがちであるが、高範囲語の順位では 116 位（言説数：51）として全言説数の約 19.9%でしか表現されていないので、必ずしもそうではない[注22)]。

表 4 グループ毎の高頻度語・広範囲語(上位 10 位)

	単棟グループ			2棟グループ			3棟以上グループ		
	Rank	キーワード	頻度	Rank	キーワード	頻度	Rank	キーワード	頻度
高頻度語上位10	1	空間	650	1	空間	288	1	空間	148
	2	建築	522	2	住宅	188	2	住戸	99
	3	住戸	422	3	住戸	151	3	都市	91
	4	住宅	405	4	敷地	145	4	集合住宅	81
	5	建物	378	5	建物	139	5	住宅	74
	6	敷地	339	6	都市	135	6	建築	70
	7	都市	322	7	建築	132	7	中庭	66
	8	集合住宅	315		計画	132	8	敷地	56
	9	計画	251	9	環境	104	9	建物	51
	10	壁	237	10	壁	102	10	環境	50
	Rank	キーワード	言説数	Rank	キーワード	言説数	Rank	キーワード	言説数
高範囲語上位10	1	空間	128	1	敷地	49	1	空間	26
		敷地	128	2	空間	46	2	敷地	24
	3	建物	119	3	住宅	45	3	住戸	23
	4	住宅	105	4	建物	41	4	建物	20
	5	建築	99	5	計画	39		集合住宅	20
	6	集合住宅	97	6	構成	36	6	建築	19
	7	計画	94	7	環境	34		構成	19
		構成	94		建築	34		住宅	19
		住戸	94	9	集合住宅	33	9	計画	16
	10	壁	93		住戸	33		壁	16

また、設計者は脈絡性を連想させる「周辺」「関係」「連続」「開放」（左から高頻度 39・25・51・86 位）といった用語もしばしば使用しているのが分かる。

グループ毎の高頻度語及び広範囲語の上位 10 位までのキーワードも 3 棟以上グループの「中庭」以外はグループの間に目立つキーワードは見られない（表 4）。

６．住棟数別グループ毎の言説の特徴分析

ここからは、住棟数によって言説がどう現れるのかを、偏って多く出現するキーワードをそのグループの特徴語と見做して抽出分析する。但し、グループ毎の特徴語は、言説数が違うことから、キーワード出現の偏り具合について特徴度という概念を用いて抽出する[文9)注23)]。

表 5 は、グループ毎に高い特徴度を持つ上位 50 位までのキーワードリストの一部である。上位 50 位までのみではあるが、単棟グループと他二つのグループとの間には、両方で上位にランクしているキーワードが皆無である一方、2 棟グループと 3 棟以上グループとの間には、順位は違うものの 9 語[注24)]のキーワードが共に上位にランクしている。2 棟グループと 3 棟以上グループの言説の特徴が単棟グループの言説よりは少し近い関係にあるといえよう。

表 5 グループ毎の上位 50 位までの特徴語の一部

単棟グループ特徴度上位50キーワード		
Rank	キーワード	特徴度
1	ファサード	1.406
2	ビル	1.331
3	形状	1.322
4	眺望	1.262
5	変更	1.262
6	下	1.260
7	階	1.259
8	水回り	1.257
9	鉄骨	1.241
10	強調	1.239
11	通り	1.228
12	窓	1.218
13	色	1.202
14	単純	1.191
15	キッチン	1.180
44	[illegible]	1.172
43	建て主	1.125
44	スラブ	1.123
45	動線	1.123
46	屋根	1.120
47	効果	1.116
48	外壁	1.116
49	合理	1.116
50	表現	1.116

2棟グループ特徴度上位50キーワード		
Rank	キーワード	特徴度
1	壁面	1.689
2	垂直	1.667
3	作業	1.656
4	最小限	1.590
5	ブロック	1.546
6	境界	1.538
7	広さ	1.529
8	感覚	1.520
9	屋上	1.514
10	通常	1.479
11	一体	1.476
12	関静	1.472
13	用意	1.457
14	プライバシー	[illegible]
15	理由	1.268
[illegible]	[illegible]	1.267
43	重要	1.265
44	場所	1.260
45	共有	1.251
46	開発	1.249
47	中庭	1.247
48	タイル	1.242
49	アクセス	1.242
50	十分	1.228

3棟以上グループ特徴度上位50キーワード		
Rank	キーワード	特徴度
1	完成	3.074
2	領域	2.447
3	形	2.306
4	中庭	2.170
5	吹抜け	2.128
6	新た	2.108
7	スケール	2.070
8	ブロック	2.049
9	考慮	2.049
10	制約	1.976
11	形成	1.976
12	共有	1.879
[illegible]	[illegible]	1.844
40	工事	[illegible]
41	発生	1.476
42	道	1.476
43	㎡	1.456
44	断面	1.456
45	手	1.447
46	設備	1.447
47	私的	1.441
48	独立	1.441
49	時代	1.441
50	風	1.419

グループ毎の特徴語の傾向を比較すると、

・<u>単棟グループ</u>では、「ファサード（1 位）」「形状（3 位）」「水回り（8 位）」「鉄骨（9 位）」「窓（12 位）」「キッチン（15 位）」など、他グループと比べ、オブジェとしての住棟扱いや住戸内部の性能、または設備・構造・材料などと関連したキーワードが特徴語として得られている。

・<u>2 棟グループ</u>では、「壁面（1 位）」「ブロック（5 位）」「屋上（9 位）」「庭（30 位）」「共有（45 位）」「中庭（47 位）」など、一概には括れないキーワードが特徴語として得られたが、単棟グループには上位にランクしてない外部空間と関連したキーワードが少し増えたのが分かる。

・<u>3 棟以上グループ</u>では、「中庭」が特徴語（4 位）として上位にランクしているのが他グループとの大きな差である[注25)]。「中庭」と共に、「領域（2 位）」「スケール（7 位）」「共有（12 位）」「開放（22 位）」など、他グループの特徴語と比べ、公共性や外部空間との関係を連想させるキーワードが多く見られる。

なお、図 3 はそこから抜粋した一部のキーワード群の特徴度をグループ間、比較してみたグラフである。厳密な選定ではないが、群棟性を示す「中庭」「領域」や単棟グループの特徴語である「ファサード」など 15 語とした。

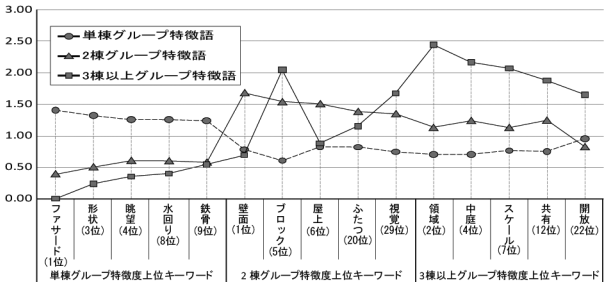


図 3 各グループの特徴語の特徴度比較

７．対応分析による各グループと特徴語の関係の視覚化

対応分析とは、クロス表を対象に行と列を調整して、相関の強いものどうしが隣接するように並び替えを行うデータマイニングの手法の一種である[文10)]。

ここからは、前章から得たキーワード（特徴語、列）と住棟数別グループ（行）との関係を該当キーワードが出現する言説の数をもとに、対応分析を行い、視覚化する。

図 4 は対応分析の結果である[注26)]。布置図を見ると、

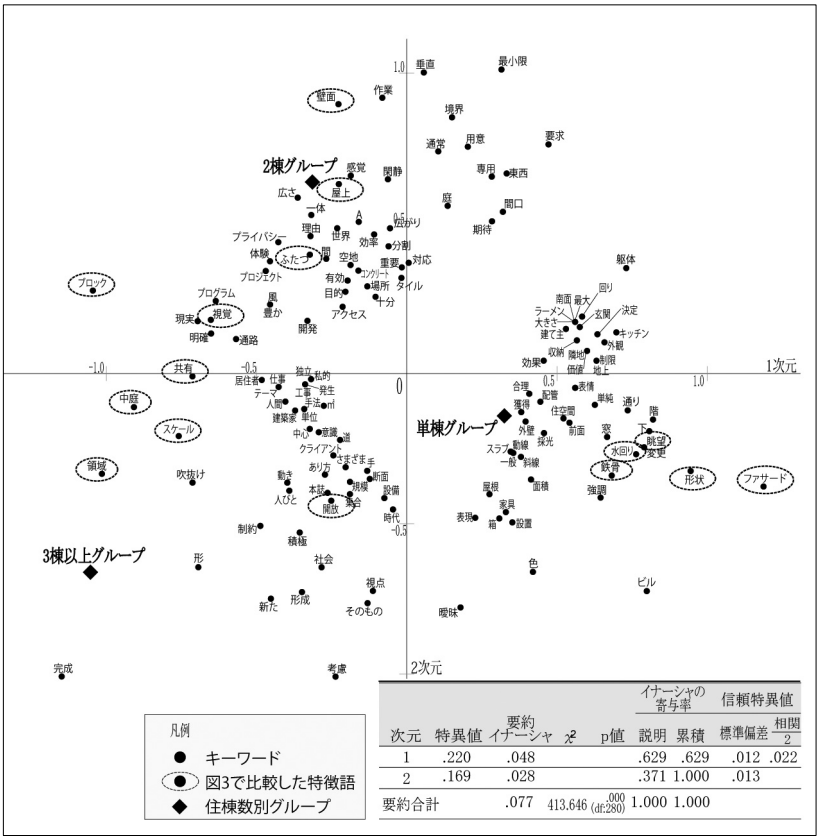
・<u>3 棟以上グループ</u>は布置図の左下側に布置しており、付近に「領域」「スケール」「中庭」や「吹抜け」「完成」「形」などが布置しており、外部空間と関係の高いキーワードが 3 棟以上グループの言説の特徴であると推測される。

・<u>単棟グループ</u>は布置図の右側に布置しており、「ファサード」「形状」や「外壁」「スラブ」「屋根」「採光」などが近い位置にあることから、住棟自体の表現と密接したキーワードが多いのが単棟グループ言説の特徴であるといえよう。

・<u>2 棟グループ</u>は左上側に布置しており、付近にあるキーワード群を見ると、他グループと比べてこれといった特徴は見えないが、「庭」「分割」「ふたつ」「空地」「通路」「ブロック」など、3 棟以上グループよりは少ないが

外部空間に関するキーワードがあり、群棟のデザインと関係のありそうなキーワードが多数布置している。

以上、対応分析を通じて住棟数別グループとグループ毎の特徴度上位 50 位までのキーワード(特徴語、計 141 語)との関係の視覚化を試み、前章の特徴分析の結果の理解を深めた。



次元	特異値	要約 イナーシャ	χ^2	p値	イナーシャの寄与率 説明	累積	信頼特異値 標準偏差	相関 2
1	.220	.048			.629	.629	.012	.022
2	.169	.028			.371	1.000	.013	
要約合計		.077	413.646 (df:280)	.000	1.000	1.000		

図 4 対応分析結果

8．4 棟以上の群棟型集合住宅の設計者言説の特徴

表 6 は 4 棟以上の群棟型集合住宅の言説から抽出した上位 10 位までの高頻度語と広範囲語のリストである。

3 棟以上グループの高頻度語と広範囲語のリスト(表 4)とほぼ同じキーワード群がランクしているが分かる。

また、表 7 は三つ以上の言説で出現するキーワードを対象に住棟数毎の特徴度を求めた上位 10 位までのリストである[注 27)]。サンプル数が少ないため一概には言えないが、住棟数に関係せず共有空間に関するキーワードが上位にランクしている。住棟数が 5 棟以上の場合、「中庭」が順位外であることから、「中庭」は住棟数が 3 棟、4 棟の群棟型集合住宅の言説において最も語られていると思われる。

表 6 4 棟以上群棟型集合住宅の高頻度語・広範囲語(上位 10 位)

高頻度語上位10 Rank	キーワード	頻度	広範囲語上位10 Rank	キーワード	言説数
1	空間	33	1	空間	7
2	都市	32		敷地	7
3	建築	22	3	建物	6
	集合住宅	22		住宅	6
5	中庭	19	5	外部	5
6	敷地	18		建築	5
7	住戸	15		集合住宅	5
	住宅	15		住戸	5
9	道路	12		都市	5
10	設計・部分	11		配置	5

表 7 4 棟以上群棟型集合住宅の特徴語

住棟数4棟の言説 特徴度上位10 Rank	キーワード	頻度	住棟数5棟の言説 特徴度上位10 Rank	キーワード	特徴度	住棟数6棟の言説 特徴度上位10 Rank	キーワード	特徴度	住棟数7棟の言説 特徴度上位10 Rank	キーワード	特徴度
1	中庭	1.756	1	構成	1.989	1	スケール	4.111	1	環境	2.126
2	テラス	1.561		社会	1.989		企画	4.111		周辺	2.126
	パーゴラ	1.561		設計	1.989		廊下	4.111		共有	2.126
	可能	1.561		道路	1.989		マンション	4.111		個室	2.126
	確保	1.561		変化	1.989		周辺	4.111		領域	2.126
	構造	1.561		さまざま	1.989		共有	4.111		タイプ	2.126
	住まい	1.561	7	部分	1.492		個室	4.111		位置	2.126
	低層	1.561		機能	1.492	8	生活	3.083		延長	2.126
	壁	1.561	9	都市	1.194		ユニット	3.083		関係	2.126
	領域・他	1.561		外部・他	1.194	10	外部・他	2.467		街・他	2.126

9．結論

以上、建築専門誌に掲載された、日本の小規模都市型集合住宅に対する設計者言説を対象に、テキストマイニングにより、その特徴を考察した。とりわけ、複数住棟からなる群棟型集合住宅の言説の特徴分析を試みた。

その結果、住棟数別グループ毎の特徴語の出現傾向の差が明らかになり、住棟数が多くなるにつれて外部空間や周辺環境と関連した特徴語が増えるのが分かった。

調査した事例の年度からすれば、群棟型集合住宅が発生し始めた時期を対象にしており、外部・環境などに関る計画の特徴を形成し、「群棟型」の特徴をメディアで発信してきたといえる。既に 30 年ほど前から日本の都市住宅の設計者たちは、どうすれば周りの住宅環境と共存でき、かつ住宅環境へ貢献できる都市住宅が作れるかを悩みつつ、一つの解決策として「群棟型」住棟計画を提案してきたのではないかと思われる[注 28)]。

すなわち「群棟型」という概念は、設計者の言説からは「中庭」のように用語としては統一的に意識されていないものの、これからの既成市街地の個別更新による住宅計画の際、社会性あるいは公共性を持つ都市住宅のあり方の一つとして、「群棟型集合住宅」の持つ可能性を示唆していると思われる。

本論の事例は、おおむね建築基準法上の 1 敷地に、容積率制限を活用した小規模な建築計画が多く、住宅供給とその需要に応える群棟型のあり方を浮かび上がらせることになり、市街地における現実的な集合住宅計画の典型の特徴を概略的・間接的に窺がうことができた。しかし、作品の具体的な計画意図まで精査するには、テキ

ストマイニングの限界もあり、さらに研究を深めることが必要であることを実感した。

最後に、テキストマイニングは、未だに発展の途上にあるのが実状であり、また、「自然言語の曖昧性」「ツールの誤謬」「得られた情報の汎用性の傾向」等、新しい知見を得るには幾つかの限界が指摘されているのも否定できない。しかし、これからもその技術の発展が期待されており、今後、設計者の言語活動と設計活動との関係を明らかにする有益な研究方法の一つとして貢献できると思われる。

注

注 1) 本論では、前稿[文3)]の「分棟型」の棟を分けるというニュアンスでなく、複数棟が群をなしているという点で「群棟型」と名称を変更した。

注 2) 文章データの有効活用を目的としているデータマイニングの一種。

注 3) 「新建築」誌に掲載した、建築家の言説の中から、建築家の住宅論・都市論・創作論等について、kj 法を用いて明らかにした研究。

注 4) 絶対頻度の高いキーワード。

注 5) 広い範囲のテキストで使われるキーワード。

注 6) 形態素解析(Morohlogical analysis)とは、連続した一連の文字列を文法的に意味のある単位の構成要素に分割し、各要素の文法的素性(品詞など)を決定することである。

注 7) 人間が日常使う自然言語を計算機によって処理すること。

注 8) 2007.10.現在消失した事例と、調査時に敷地の分割状況が不明確であった森山邸(設計：西沢立衛)は除外。2013.5.現在、4 件がさらに消失したが分析に含めた。収集事例は総 272 件で、その中 95 事例を 2 棟以上の群棟型として分類した。

注 9) 4 棟以上の群棟型集合住宅は 3 棟以上グループに含めて分析する。

注 10) 2 事例以上が 1 言説で説明されているのが理由である。また、言説は本文記事を中心とした。事例に関わる数人の関係者の言説は一つにまとめて扱うことにし、対談文も含めた。

注 11) 敷地は、雑誌で表明された敷地とする。なお、雑誌における敷地は、多くは建築基準法上の敷地である。かならずしもそうではない少数の事例を含んでいるが、雑誌に表明された敷地は、調査・変更を行わず、そのまま利用し手続きを一定にした。
平均建蔽率・容積率などは、表記されたデータを優先し、無い場合は、建蔽率=建築面積/敷地面積、容積率=延床面積/敷地面積で計算した。平均住戸数は図面から読み取ったものであり、他用途は除外した。
群棟の分類では、図 2 の下部の例に示すように、1 建築として見せるために付加された小規模な付加物により住棟を結ぶ事例の場合、逆に棟が建築計画的に連続していないと判断し、複数の棟として分類作業を行った（詳細は、文 3 参照）。

注 12) 京都大学情報学研究科－NTT コミュニケーション科学基礎研究所共同研究で開発されたオープンソース形態素解析エンジン。

注 13) 形態素解析の際、調査単位(単位の長さや同形異議語の判別等についての単位の幅)の設定などが問題となるが、本論では Mecab の基本性能(Mecab は、IPA 辞書(ver2.7.0)の利用を推奨しており、本辞書では、情報処理振興事業協会(IPA)で設定された IPA 品詞体系(THiMCO97)に基づいて、一部修正や改良が行われている。)に従う。

注 14) Mecab は未知語等が追加できる使用者辞書の追加ができる。使用者辞書には約 1,490 余の単語を登録した。

注 15) 「送り仮名の統一」や「漢字・略語の統一」、「誤字等、原文の単語の内容の書き換え」等。

注 16) 名詞の中、品詞細分類 1 の「サ変接続」、「ナイ形容詞語幹」、「形容動詞語幹」、「一般」、「固有名詞」を抽出する。一般的には動詞、形容詞、名詞の三つの品詞種類が文の重要語とされるが、本論では名詞のみを調査対象とする。入手先　URI: http://www.unixuser.org/ /~euske/doc/postag/index.html#chasen(アクセス日:2013.2.)

注 17) 延べ語数：個々の自立語が同じ単語かどうかにかかわらず、どの単語も同じように一つ二つと数えた総数。

注 18) 異なり語数：違う単語の種類を数えた総数。

注 19) 厳密には、「見出し語」と言うべきが、調査対象が名詞のみである等から、キーワードと呼ぶことにする。

注 20) 原因の一つとしては、作品名などの固有名詞の使用にあると思われるが、普通名詞も多く、一概には言えない。

注 21) 「街区」は、「街(名詞・一般)」＋「区(名詞・接尾)」として分析された。全文からの「街区」の頻度は 25 であり、その他「地区」(頻度：15)、「住区」(頻度：10)、「区域」(頻度：5)などが見られた。

注 22) 「家」の場合も高頻度(19 位)と広範囲語(87 位)との、両方の順位の差が大きいが、一部の言説で偏って多用されたからではないかと思われる。

注 23) 特徴度は以下の式①から計算する。
特徴度＝キーワードのグループからの出現言説数の実測値/出現言説数の期待値 ……①
[出現言説数の期待値＝(キーワードの全グループからの出現言説数×全キーワードのグループからの出現言説数)/全キーワードの全グループからの出現言説数]。但し、全てのキーワードの特徴度をグループ毎に求めると、同じ特徴度の持つ特徴語が多くなり分別力が薄くなるため、全言説数の約 10%である 25 個以上の言説で出現するキーワードのみを対象にグループ毎の特徴度を求める。

注 24) 「中庭」「ブロック」「共有」「現実」「明確」「視覚」「通路」「プログラム」「風」

注 25) 「中庭」は「コート」を含め、頻度・出現言説数の両方共に、3 棟以上グループの方が相対的に高く、その意味合いも、他グループでは住棟の小振り操作による小さな外部空間づくりに関する表現が多い半面、3 棟以上グループでは纏まった共用の外部空間づくりに関する表現が多かった。

注 26) SPSS(18.0)で分析した。1 次元が約 63%、2 次元が約 35%の割合で元の情報を説明している。又、χ^2検定(自由度 280、P 値=0.000)結果、行列は関係があるといえる。行ポイントと列ポイントの概要は割愛するが、行ポイントの場合、1 次元では 3 棟以上グループ、2 次元では 2 棟グループの寄与率が高かく、列ポイントの場合、1 次元では「領域」「ファサード」「中庭」「ブロック」順に、2 次元では「壁面」「垂直」「作業」順に高かい寄与率の結果を見せた。

注 27) 4 棟以上群棟型集合住宅の言説からは、建物と外部空間を同等に重要視しながら、周辺環境のスケールを尊重し、且つ街や都市に開かれた開放的な住居環境づくりを強調する表現が多かった。

注 28) 前稿[文3)]の研究で、群棟型として分類された 95 事例は全期間にわたってほぼ均等に分布していた。

参考文献

文 1) 京都造形芸術大学篇: 住まいの解剖学, 角川書店, 1998

文 2) 服部岑生, 鈴木雅之, 泉幸甫: 住宅地の更新における小規模集合住宅群の建築計画-住宅地更新のための住宅形式に関する研究-, 日本建築学会住宅系研究論文報告会論文集 1, pp.299-308, 2006.

文 3) 李尚勲: 現代日本における小規模都市型集合住宅の計画特性に関する研究-住棟の分棟類型に関する基礎的研究-, 大韓建築学会計画系論文集, v.24 n.10, pp.47-58, 2008.10.

文 4) 奥山信一, 坂本一成；戦後「新建築」誌にみられた建築家の住宅観-建築家の住宅論に関する研究-, 日本建築学会計画系論文報告集, 第 428 号, pp.125-135, 1991.10

文 5) 奥山信一, 斉藤千尋, 坂本一成；戦後「新建築」誌にみられた建築家の都市観-建築家の住宅論・都市論に関する研究-, 日本建築学会計画系論文報告集, 第 444 号, pp.49-59, 1993.2

文 6) 奥山信一, 持田英明, 坂本一成；戦後「新建築」誌にみられた建築家の創作の主題-建築家の創作論に関する研究-, 日本建築学会計画系論文報告集, 第 454 号, pp.77-86, 1993.12

文 7) 小林祐司, 寺田充伸, 佐藤誠治:テキストマイニングを活用したアンケートにおける自由回答の分析と生活環境評価, 日本建築学会計画系論文報告集, 第 671 号, pp.85-93, 2012.1

文 8) 夏目欣昇, 若山滋, 堀桂輔:評論における名詞と修飾語の出現傾向の考察-西洋近代建築評論にみられる単語関係の可視化その 1 -, 日本建築学会計画系論文報告集, 第 632 号, pp.249-256, 2008.10

文 9) 上田太一郎監修: 事例で学ぶテキストマイニング, 共立出版, 2008

文 10) 石田基広: R によるテキストマイニング入門, 森北出版, 2008