

1. 到達容易性による評価

そこで、[11]と同様に、ある語を入力した場合に、その関連語を関連度の高いものから順に提示することを考え、先頭の語からはじめて、一つ下位の語をみる確率を p とし、 n 番目の語に到達する確率を p^{n-1} 、ある語の関連語を見に行く確率を q (ただし、 $p+q<1$) として、それぞれの資料群設定における到達容易性の計算を行った。 $p=0.89$, $q=p/20$ として、関連語は第 60 位まで、別の関連語を辿る回数は 4 回までに制限して、「土器」から「石鏃」に至る到達可能な全経路の計算を行い、全経路の確率の和を求めたところ、

分割 1 (コレクションをそのまま資料群とした場合) : 0.168589930619463

分割 2 (階層構造で資料群を分割した場合) : 0.047174548107095

分割 3 (人手で分割位置を設定した場合) : 0.0744339146642573

分割 4 (階層構造を反映した計数を行った場合) : 0.0593648767031672

となった。

専門家によって、「石鏃」は「土器」の関連語としてはふさわしくないと判定されているため、「土器」から「石鏃」へ至る到達可能性は低い方が望ましいとすると、望ましい分割は順に、分割 2 > 分割 4 > 分割 3 > 分割 1 となる。

次に、「土器」の関連語としては比較的下位で、専門家が関連語としてふさわしいと判定しているものの中から「高杯」を選び、「土器」から「高杯」に至る到達可能な全経路の計算を行い、全経路の確率の和を求めたところ、

分割 1 (コレクションをそのまま資料群とした場合) : 0.00049785187787727

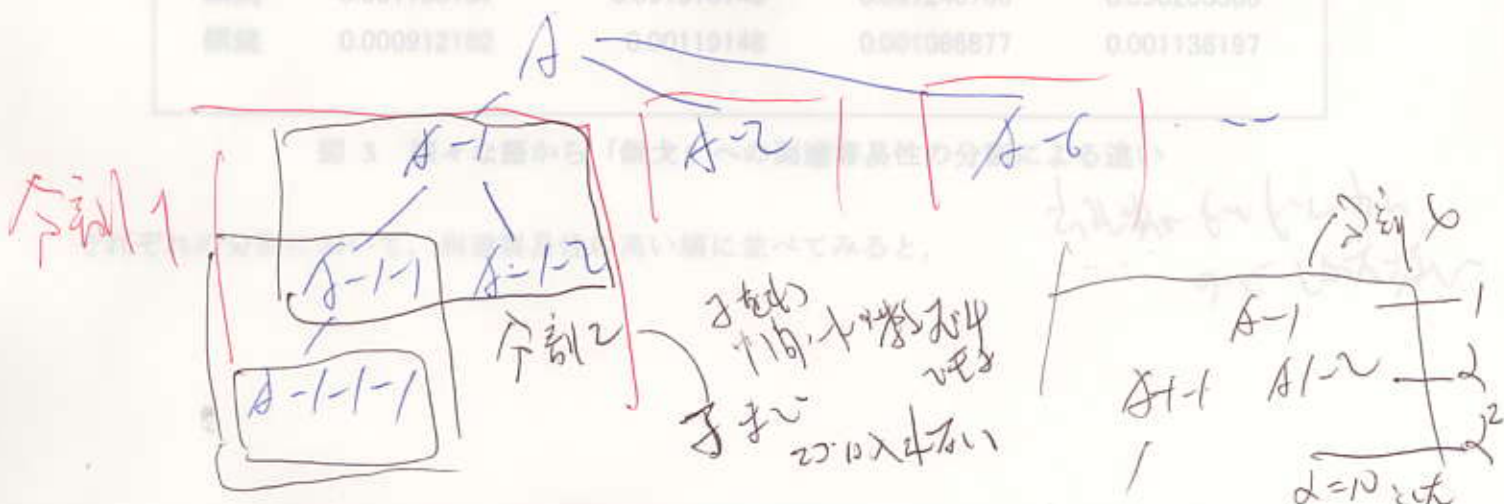
分割 2 (階層構造で資料群を分割した場合) : 0.00142262644849791

分割 3 (人手で分割位置を設定した場合) : 0.00203064928687728

分割 4 (階層構造を反映した計数を行った場合) : 0.00425238737485268

となった。こちらは逆に、到達容易性が高いほど望ましいとすると、望ましい分割は順に、分割 4 > 分割 3 > 分割 2 > 分割 1 となる。ただし、異なる分割間で、到達容易性の絶対値で比較することには問題もある。

そこで、様々な語から出発して「高杯」に至る到達容易性を 3 種の分割のそれぞれについて求め、同じ分割の中でそれらの比較をしてみる。出発語として「石器」、「土器」、「土偶」、「埴輪」、「銅鐸」、「銅剣」、「銅鏡」を設定した場合の計算結果を図 2 に示す。



	分割 1	分割 2	分割 3	分割 4
石器	0.000194056	0.000369437	0.000549275	0.000267481
土器	0.000497852	0.001422626	0.002030649	0.004252387
土偶	0.000345532	0.000186265	0.000621021	0.001287410
埴輪	9.90E-03	0.015456021	0.028347691	0.007692792
銅鐸	0.001193007	0.000708322	0.001994376	0.000154921
銅剣	0.006404267	0.000203992	0.008341554	0.000897488
銅鏡	0.002636213	0.003337144	0.009224251	0.000296631

図 2 様々な語から「高杯」への到達容易性の分割による違い

それぞれの分割において、到達容易性の高い順に並べてみると、

分割 1 : 埴輪 > 銅剣 > 銅鏡 > 銅鐸 > 土器 > 土偶 > 石器

分割 2 : 埴輪 > 銅鏡 > 土器 > 銅鐸 > 石器 > 銅剣 > 土偶

分割 3 : 埴輪 > 銅鏡 > 銅剣 > 土器 > 銅鐸 > 土偶 > 石器

分割 4 : 埴輪 > (土器) > 土偶 > 銅剣 > 銅鏡 > (石器) > 銅鐸

となる。

これらを比べてみると、「土器」の相対的順位が最も高いのは分割 2 であるが、同時に「石器」の順位も高くなっている。専門家により「高杯」は「石器」の関連度としてはふさわしくないと判定されているため、「石器」から出発した場合の到達容易性は低い方が望ましいことを勘案すると、分割 3 のほうがよりよい分割になる。

同様に、「銅戈」に至る到達容易性を 3 種の分割のそれぞれについて求め、「石器」、「土器」、「土偶」、「埴輪」、「銅鐸」、「銅剣」、銅鏡」から出発した場合の計算結果を図 3 に示す。

	分割 1	分割 2	分割 3	分割 4
石器	3.24E-07	3.01E-07	7.13E-07	4.08E-06
土器	1.86E-07	2.29E-07	4.12E-07	2.38E-06
土偶	4.44E-07	4.55E-08	5.89E-07	5.42E-06
埴輪	4.31E-06	3.00E-05	4.88E-05	9.32E-06
銅鐸	0.102808997	0.102965398	0.104108761	0.102889490
銅剣	0.087795157	0.091310143	0.091240786	0.090203508
銅鏡	0.000912182	0.00119148	0.001086877	0.001136197

図 3 様々な語から「銅戈」への到達容易性の分割による違い

それぞれの分割において、到達容易性の高い順に並べてみると、

関連容易性 分割1: 銅鐸 > 銅剣 > 銅鏡 >> 埴輪 > 土偶 > 石器 > 土器

分割2: 銅鐸 > 銅剣 > 銅鏡 >> 埴輪 > 石器 > 土器 > 土偶

分割3: 銅鐸 > 銅剣 > 銅鏡 >> 埴輪 > 石器 > 土偶 > 土器

分割4: 銅鐸 > 銅剣 > 銅鏡 >> 埴輪 > 土偶 > 石器 > 土器

となり、3種の分割の間で相対順位の差異はほとんど見られない。実際には、いずれの分割においても「埴輪」以下からはほとんど到達する可能性はないと見るべきである。もともと「銅戈」は出現数が少なく、関連語の数も少ないため、このような語に対して分割を変えることの効果は見られなかった。

分割1 (コレクションをそのまま資料群とした場合): 0.168589930619463

分割2 (階層構造で資料群を分割した場合): 0.047174548107095

分割3 (入手で分割位置を設定した場合): 0.0744339146642573

分割4 (階層構造を反映した計算を行った場合): 0.0593648767031672

となった。

専門家によって、「石器」は「土器」の関連語としてはふさわしくないと判定されているため、「土器」から「石器」へ至る到達可能性は低い方が望ましいとする。望ましい分割は順に、分割2 > 分割4 > 分割3 > 分割1となる。

次に、「土器」の関連語としては比較的単純で、専門家が関連語としてふさわしいと判定しているものの中から「高杯」を選び、「土器」から「高杯」に至る到達可能な全経路の計算を行い、全経路の確率の和を求めたところ、

分割1 (コレクションをそのまま資料群とした場合): 0.00049785187787727

分割2 (階層構造で資料群を分割した場合): 0.06142262644849791

分割3 (入手で分割位置を設定した場合): 0.00203064928687728

分割4 (階層構造を反映した計算を行った場合): 0.00425238737485248

となった。こちらは逆に、到達容易性が高いほど望ましいとすると、望ましい分割は順に、分割4 > 分割3 > 分割2 > 分割1となる。ただし、異なる分割間で、到達容易性の絶対値で比較することには問題もある。

そこで、選んだ語から出発して「高杯」に至る到達容易性を3種の分割のそれぞれについて求め、選んだ分割の中でそれらの比較を試みる。出発語として「石器」、「土器」、「土偶」、「埴輪」、「銅鐸」、「銅剣」、「銅鏡」を設定した場合の計算結果を図2に示す。

