

養護教諭を目標として学ぶ学生に対する 水質調査の教育的効果

千 田 眞喜子

本研究では、水質調査演習の教育効果の確認を目的として、演習レポートの考察・感想の自由記述をテキストマイニングにより分析し、学生の水質調査の意義の理解状況を探り、以下の知見が得られた。頻出語は「水」、「川」、「綺麗」、「汚れ」、「水道水」、「多い」、「考える」、「水中」、「高い」の順に多く、これらの語に関する記述が多かった。また、「水質調べの重要性の把握」、「試薬により透明な水の色が変化して水質が確認できること」、「水の安全性」、「塩素消毒」についての記述が多かった。演習を通じて水質に直に触れその安全性に対する認識を深めていることがわかる。座学で知識として学ぶだけでなく、水質調査を演習に取り入れることにより、養護教諭を目指す学生の水質への理解が深まり教育効果が上がることが確認できた。

キーワード：共起ネットワーク、教育的効果、ミニマムスパンニングツリー、テキストマイニング、水質調査

This study was conducted to ascertain educational effects of water quality surveys for students to be nursing teachers. The discussion parts of the students' reports were analyzed by a method of text mining to grasp their understanding level on the importance of the surveys. Results show that frequently appearing words were "water," "river," "beautiful," "dirty," "tap water," "many," "thinking," "in water," "high" in descending order, and these words were included in numerous descriptions. Many descriptions were related to "grasping the importance of water quality survey," "being able to check the water quality by changing the color of transparent water by reagent", "water safety," and "chlorine disinfection." It is indicated that the students have deepened their recognition on the water quality and its safety through the survey exercise. It is confirmed that incorporating the water quality survey into the curriculum increases educational effects and enhances the understanding of the water quality.

Keywords : co-occurrence network, educational effect, minimum spanning tree, text mining, water quality survey

I. はじめに

養護教諭には、「学校環境衛生」に関して、1) 学校薬剤師が行う検査の準備、実施、事後措置に対する協力、2) 教職員による日常の学校環境衛生活動への協力・助言等の職務がある（電子政府の総合窓口、2018；文部科学省、2018）。学校環境衛生活動の日常点検の中に、飲料水・雑用水の水質及び飲料水等の施設・設備点検があり、定期

検査では飲料水等の水質及び施設・設備の検査があり、臨時検査もある。現場では、実際に養護教諭が水道水の残留塩素等の計測に関わることもある。このように、「水質調査の意義」を養護実習指導において学ぶことは重要である。そこで、本研究では、水質調査演習の教育効果の確認を目的とし、学生の水質調査の意義の理解状況を分析・検討した。

II. 分析手法

本研究の対象者は、京都市の A 大学の、養護教諭を目指して養護実習指導の科目を履修する 3 年生 14 名である。2018 年 1 月の水質調査演習において、各自が興味ある水（ミネラルウォーター、雪、河川水、雨水、水道水）の残留塩素濃度（(株) 共立理化学研究所製、型式 WAK-CIO/DP で計測）と、川の水調査セット（(株) 共立理化学研究所製、型式 TZ-RW で計測）の項目（COD・アンモニウム態窒素・亜硝酸態窒素・硝酸態窒素・リン酸態リン）の濃度について計測した。その結果を水質調査に関するレポートとして Microsoft Word で提出させ、その自由記述の考察・感想の文章のテキストデータを分析に用いた。

自由記述の文章にテキストマイニング（松村・三浦、2009）を使用し、統計処理は KH coder（KH coder、2018；樋口、2014）を用いた。テキストマイニングとは、“定性的な情報のテキストデータ”を計算機で“定量的に解析”して有用な情報を抽出するための様々な方法の総称である（松村・三浦、2009）。テキストマイニングでは、一般的に文章を単語に分割し、各単語の品詞を求める形態素解析を行い、頻出語を類義語や同一の概念でまとめるコーディングと呼ばれる作業を行う。動詞や形容詞など活用のある語を抽出する際、それらの語は基本形に直して抽出する（樋口、2014）。ただし、助詞・助動詞をはじめ、どのような文の中でも出現するであろう品詞は分析から外した。また、平仮名のみからなる語についても分析時に利用しにくい一般的な語なので分析

時に除外した（樋口、2014）。

自由記述の文章を形態素分析により一語一語に分け、語の出現回数を調べた。形態素解析器は茶筌（奈良先端科学技術大学院大学情報科学研究科自然言語処理学講座松本研究室、2018）を用いた。前処理を行った後、データベースの統計結果（総抽出語数、異なり語数）を確認し、次に抽出語リストを出力した。

分析には共起ネットワーク（最小スパニングツリー）を用い視覚的に検討した。共起ネットワークとは共起の程度が強い（出現パターンが似通った）語を線で結んでネットワークのことである（Corman、2002；Osgood、1968；鈴木、2002；鈴木、2006）。語と語の関連を探索するために、関連性の可視化により判断する分析手法の 1 つである。

語と語をつなぐ線が多くなった場合には、どの線が重要なのかの手がかりがあったほうが図を解釈しやすい場合がある。重要なものを抽出、明確化するために、本研究では“一番重要な線だけで表現する最小スパニングツリー”を作成した。

III. 結果・考察

III-1 データベースの統計結果概要

形態素解析の結果の概要を表-1 に示す。文は 314 ケース、段落は 30 段落、見出しは 5 つ（ミネラルウォーター、雪、河川水、雨水、水道水）であった。総抽出語数（使用）は 7,394 (2,976)、異なり語数は 1,174 で、解析に使用した語は 975 であった。抽出語の出現回数の平均は 3.05、抽出語の出現回数の標準偏差は 8.82 であった。

表-1 抽出語の基本統計量

総抽出語数（使用）	7,394 (2,976)
異なり語数（使用）	1,174 (975)
抽出語の出現回数の平均	3.05
抽出語の出現回数の標準偏差	8.82
集計単位	文 ケース数 314
	段落 ケース数 30
見出し	ミネラルウォーター、雪、河川水、雨水、水道水

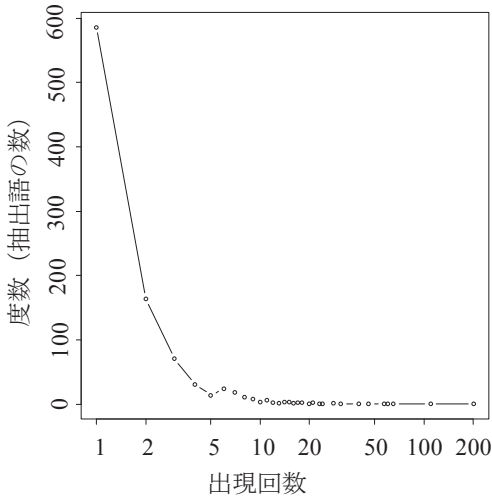


図-1 抽出語の出現回数と度数 (抽出語の数)

抽出語の出現回数別に何種類の語が用いられているかを見るために、図-1 にそれぞれの語のデータ全体における出現回数の分布を示す。横軸は抽出語の出現回数を対数軸で表し、縦軸は抽出語の度数 (語の種類) を表した。抽出語の出現回数 10 回程度から、抽出語の度数が少なくなっていることがわかった。

III-2 言葉の出現頻度

頻出上位 150 語を表-2 に示す。「水」が 110 回で最も多く、次いで「川」が 65 回、「綺麗」が 46 回、「汚れ」が 31 回、「水道水」が 28 回、「多い」が 23 回、「考える」・「水中」が 21 回、「高い」が 20 回であった。これらの語の記述が多い

ことがわかった。

次に、最も多く使われていた「水」と関連する語の検索結果を表-3 に示す。共起尺度として主なものは、共起頻度、Jaccard 係数 (村田・大沢、2017; Kautz, H., et al., 1997) があり、ここではその 2 つの尺度を示す。

「X」と「Y」の単独での出現数を $|X|$ 、 $|Y|$ 、どちらか一方が出現した回数を $|X \cup Y|$ 、両方が出現した回数を $|X \cap Y|$ とする (図-2)。

「共起頻度」とは、共起の回数であり、 $|X \cap Y|$ で計算される。

「Jaccard 係数」とは、2 つの集合間の類似性を表す指標である。どちらかが出現したうち、何回同時に出現するかで、 $|X \cap Y|/|X \cup Y|$ で計算される。X と Y の Jaccard 係数は以下の式で表される。

$$\text{Jaccard}(X, Y) = \frac{|X \cap Y|}{|X \cup Y|} \quad (1)$$

例えば、 $X=\{a,b,c\}$ 、 $Y=\{a,b,d,e\}$ のとき、 $|X \cap Y|=2$ 、 $|X \cup Y|=5$ なので、Jaccard 係数は $2/5$ になる。Jaccard 係数が大きいほど (1 に近いほど) X と Y は似ていると言える。

他にも、Simpson 係数、コサイン距離があるが、本稿では Jaccard 係数を中心に考察する。

「水」と最も共起頻度が高い語は「川」であり、Jaccard 係数が高い語は「COD」であった。また、「考える」、「思う」が共起頻度及び Jaccard 係数とも 3、4 番目に高かった。これらのことから、学生は「水」に対して川や COD について考えいろいろ思っていることを記述していた。

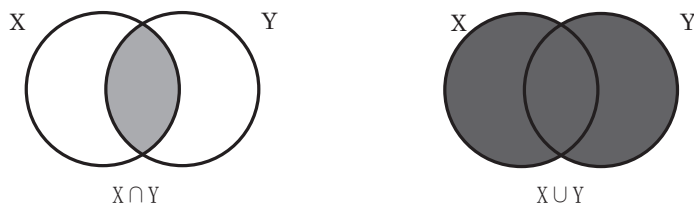


図-2 $|X \cap Y|$ と $|X \cup Y|$

表-2 頻出上位 150 語

抽出語	出現回数	抽出語	出現回数	抽出語	出現回数
水	110	水質	8	定める	5
川	65	雪	8	日本	5
綺麗	46	分解	8	問題	5
汚れ	31	飲める	7	ナチュラルウォーター	4
水道水	28	汚れる	7	パックテスト	4
多い	23	見る	7	飲料水	4
考える	21	酸態	7	汚染	4
水中	21	出来る	7	河川	4
高い	20	消費	7	学ぶ	4
結果	18	場合	7	学校	4
思う	18	場所	7	魚	4
反応	18	状態	7	驚く	4
COD	17	食べ物	7	検査	4
流れ込む	17	生物	7	見える	4
感じる	16	窒素	7	工場排水	4
分かる	16	比較的	7	今	4
ミネラルウォーター	15	予想	7	使う	4
残留塩素	15	量	7	死亡率	4
値	15	安心	6	私たち	4
調べる	15	安全	6	浄化能力	4
塩素	14	塩素消毒	6	水系感染症	4
酸素	14	加える	6	水質調査	4
低い	14	形	6	棲む	4
有機物	14	国	6	生態系	4
飲む	13	自然	6	測定範囲	4
少し	13	自分	6	通常	4
アンモニウム態窒素	12	失う	6	乳幼児	4
今回	12	植物	6	不足	4
探る	12	色	6	すべて	3
可能性	11	人々	6	ほか	3
近く	11	水質検査	6	カルキ臭	3
硝酸態窒素	11	生活	6	タンパク質	3
入る	11	測定	6	バランス	3
変化	11	存在	6	プール	3
亜硝酸態窒素	10	調査	6	リン酸態リン	3
原因	10	土壌	6	違い	3
大きい	10	道路	6	違う	3
濃度	10	判断	6	一つ	3
雨水	9	必要	6	塩素濃度	3
湖	9	流れる	6	汚い	3
出る	9	影響	5	家庭	3
上流	9	酸化剤	5	河川水	3
数値	9	指標	5	改めて	3
生活排水	9	時間	5	外	3
肥料	9	示す	5	割合	3
確認	8	十分	5	環境	3
含む	8	人間	5	関わる	3
気	8	性質	5	基準	3
持つ	8	整備	5	起きる	3
少ない	8	大切	5	技術	3

表-3 水と関連する語の検索結果

n	抽出語	品詞	全体	共起	Jaccard 係数
1	COD	未知語	12 (0.211)	12 (0.632)	0.632
2	川	名詞 C	15 (0.263)	13 (0.684)	0.619
3	考える	動詞	13 (0.228)	11 (0.579)	0.524
4	思う	動詞	11 (0.193)	10 (0.526)	0.500
5	出る	動詞	8 (0.140)	8 (0.421)	0.421
6	値	名詞 C	8 (0.140)	8 (0.421)	0.421
7	調べる	動詞	8 (0.140)	8 (0.421)	0.421
8	きれい	形容動詞	9 (0.158)	8 (0.421)	0.400
9	感じる	動詞	9 (0.158)	8 (0.421)	0.400
10	高い	形容詞	10 (0.175)	8 (0.421)	0.381
11	綺麗	形容動詞	10 (0.175)	8 (0.421)	0.381
12	今回	副詞可能	7 (0.123)	7 (0.368)	0.368
13	有機物	名詞	7 (0.123)	7 (0.368)	0.368
14	結果	副詞可能	8 (0.140)	7 (0.368)	0.350
15	数値	名詞	8 (0.140)	7 (0.368)	0.350
16	大きい	形容詞	8 (0.140)	7 (0.368)	0.350
17	反応	サ変名詞	8 (0.140)	7 (0.368)	0.350

Ⅲ-3 共起ネットワーク分析（最小スパニングツリー）による分析

頻出上位 60 語を用い、共起ネットワーク分析（最小スパニングツリー（MST））結果をもとめる。語と語を結ぶ線上の数値は Jaccard 係数で 0.166 以上を示し、出現頻度が高いほどバブルを大きく示す。

まず、語と語の共起関係を図-3 に示す。最小スパニングツリーを太線で、次数中心性は高いほど濃い灰色で示す。「有機物」が最も次数中心性が高く 10 種類の語が共起していた。次に「酸

素」・「生物」・「消費」の中心性が高く 8 種類の語が共起していた。図-3 から主にわかることは、以下のとおりである。

- 1) COD の指標と有機物の含有と生態系の相互の関わりについて記述していた。
- 2) 水質調査は大切と感じている。
- 3) 試薬の効果により水の色が変化して水質が確認できる。
- 4) 水に汚れが流れ込むこと。

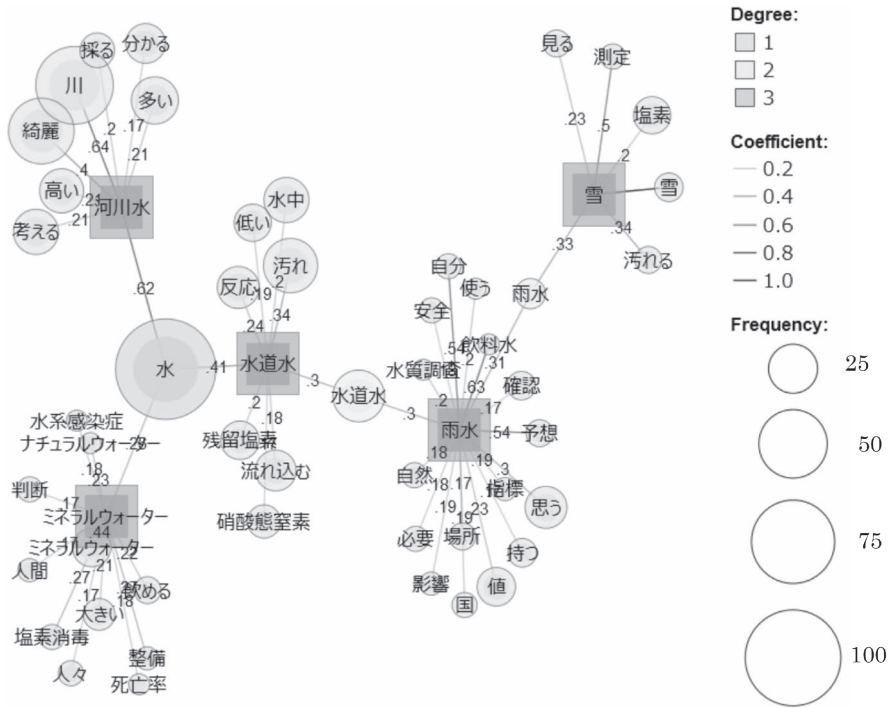


図-4 「計測した水」と「語」の共起ネットワーク（最小スパニングツリーだけ表示）。計測した水は河川水、ミネラルウォーター、水道水、雪、雨水で四角で示す。丸で表示は頻出上位60語を示す。

V. 今後の課題

水質の計測に関しては、パックテストの色の変化を標準色比色表を用いて目視で行った。比色（ひしょく）には個人差があるので、より正確な計測ができる機器を用いて計測を行いたい。また、計測前と計測後の学生の意識の変化についても今後の課題とする。

付記

本稿は、2018年9月に、水文・水資源学会2018年度総会・研究発表会（於三重大学）にて発表した原稿（千田、2018）の一部に加筆修正したものである。

引用文献：

- Corman, S. R., Kuhn, T., Mcphee, R. D., and Dooley, K. J.(2002) Studying Complex Discursive Systems Centering Resonance Analysis of Communication, *Human Communication Research*, Vol.28, No.2, pp.157-206, DOI : 10.1111/j.1468-2958.2002.tb00802.x.
- 電子政府の総合窓口 (2018) : 学校保健安全法、
http://elaws.e-gov.go.jp/search/elawsSearch/elaws_search/lsg0500/detail?lawId=333AC0000000056&openerCode=1#C、(参照：2018/03/12).
- 樋口耕一 (2014) 社会調査のための計量テキスト分析 一内容分析の継承と発展を目指して一、pp.1-233、ナカニシヤ出版。
- Kautz, H., Selman, B., and Shah, M. (1997) The Hidden Web, *AI magazine*, Vol.18, No.2, pp.27-35.
- KH coder(2018) <http://khc.sourceforge.net/>、(参照：2018/2/12).

- 文部科学省 (2018) 養護教諭の職務内容等について、
http://www.mext.go.jp/b_menu/shingi/chousa/shotou/029/shiryo/05070501/s007.htm、(参照：2018/03/12)。
- 松村真宏、三浦麻子 (2009) 人文・社会科学のためのテキストマイニング、153pp、誠信書房。
- 村田亘、大沢英一 (2017) ハイパーメディアの Jaccard 係数に着目した定義文拡張による語義曖昧性解消、日本ソフトウェア科学会第 34 回大会 (2017 年度) 講演論文集、34、pp.49-58。
- 奈良先端科学技術大学院大学情報科学研究科自然言語処理学講座 (松本研究室) (2018) 茶筌 ChaSen -- 形態素解析器、<http://chasen-legacy.osdn.jp/>、(参照：2018/2/12)。
- Osgood, C. E. (1959) The representational model and relevant research methods. In: Pool, I. de S. (Ed.), Trends in content analysis. Urbana, Ill.: University of Illinois Press, pp.33-88.
- 千田真喜子 (2018) 養護教諭を目指す学生に対する水質調査の教育的効果、水文・水資源学会 2018 年度研究発表会要旨集、pp.294-295。
- 鈴木努 (2002) ニュース・テキストのイシュー連関構造分析、マス・コミュニケーション研究, Vol.60, pp.137-152. DOI : 10.24460/mscom.60.0_137.
- 鈴木努 (2006) 二〇〇五年衆議院選挙における三大紙の社説比較：概念ネットワーク分析の適用、マス・コミュニケーション研究, Vol.69, pp.2-21. DOI : 10.24460/mscom.69.0_2.