

## 卒業時アンケート「大学への総合的満足度」の理由のテキストマイニング

Text Mining of Reasons for Overall Satisfaction with the University  
in a Graduation Survey橋本 弥生<sup>\*1</sup>Mio TSUBAKIMOTO<sup>\*1</sup><sup>\*1</sup> 東京都立大学 大学教育センター<sup>\*1</sup>University Education Center, Tokyo Metropolitan University

Email: mio@tmu.ac.jp

あらまし：学部生の卒業時アンケートにおける大学満足度の自由記述データについて、コロナ禍の時期や GPA の高低で記述内容や量に差があるかをテキストマイニングによって探索的に分析した。その結果、全年度と GPA に共通する満足理由として「教員との距離の近さ」や「サークル活動」が挙げられた。年度別では、コロナ禍の時期は満足理由のバリエーションや他の年度と共通して出現する語が少なくなっていた。GPA 別では、高群は学修活動が、低～中群は周辺的な活動が満足理由としてあげられていた。

キーワード：学生調査、自由記述、テキストマイニング、大学満足度、GPA

## 1. 背景と目的

大学の授業がコロナ禍によりオンライン中心となった 2020 年度から 2021 年度を経て、2024 年度現在では、多くの大学で対面授業が復活し、学生の生活はほぼコロナ以前に戻ってきている。この間、学生の大学に対する満足度に変化はあったのだろうか。あったとすれば、どのような変化だろうか。

本研究では、都内の総合大学である公立 A 大学が毎年 3 月に学部 4 年生に対して実施している学生調査「卒業時の学修成果等アンケート」の回答から、コロナ禍の中心的な期間（2020 年 3 月～2021 年 3 月）とその前後を含む 2017 年度から 2022 年度の 6 年間分の自由記述データを用いて、その内容と総合的満足度の評定値、および卒業時 GPA\* (GPA\*とは、評定 0 (欠席) を除き、1 から 5 の評定値のみの成績データをもとに算出した修正版 GPA である) との関係テキストマイニングによって探索的に分析する。

## 2. 方法

## 2.1 分析対象

2017 年度から 2022 年度にかけて収集された上記調査のうち、以下の全ての要素において欠損がなかったデータ計 1,946 件を分析対象とした。

- 大学への総合的満足度 (1:不満～4:満足) の評定値および評定理由の自由記述
- 卒業時の GPA\*

## 2.2 テキストデータのクレンジングと使用する品詞

KH Coder3 正式版<sup>(1)(2)</sup>に搭載された MeCab を用いて 6 年間の自由記述データに対して形態素解析を行い、48,759 語 (異なり語数 2,475 語) を得た。形態素解析の結果より、今回の分析に用いる品詞を [名詞], [サ変名詞], [動詞], [形容詞], [副詞] とした。一般的に、文章内容をとらえるために必要とされている品詞は名詞、動詞、形容詞である。今回はこれ

らに加えて、名詞の多様性をできる限り担保するために [サ変名詞] を、程度についての把握も必要と考えられたため [副詞] をそれぞれ加えた。

さらに、文脈から同じ意味をさしていると判断された語群は、それらを代表する語に表記を統一した (例: 先生, 教員, 教授 → 「教員」としてまとめた)。

## 3. 結果と考察

## 3.1 満足度と GPA\* の記述統計

年度ごとのデータ数 (N) および、総合的満足度と GPA\* の平均値と標準偏差を表 1 に示す。表 1 より、総合的満足度の平均値は 3.20 から 3.53 の範囲であり、非常に高い値に集中していることがわかる。

分析対象者の評定値の分布を各年度において確認したところ、「満足」を意味する選択肢 3 と 4 の割合の合計は毎年約 90% であった。したがって、今回分析対象とするデータの約 90% は、この調査において大学への満足度が比較的高いと回答していたことが明らかになった。さらに、同様に卒業時 GPA\* の分布を確認したところ、2 未満の学生は 7%～17% であり、対象者の 80% 以上は成績が中程度から高い学生であった。以上より、分析対象者の傾向の偏りに留意しつつ結果を解釈する必要があることが示された。

なお、これ以降のテキストマイニングでは、卒業時 GPA\* の値を「2 未満」「2 以上 3 未満」「3 以上」の 3 カテゴリに分類して用いる。

表 1 年度ごとの記述統計

年度	総合的満足度		卒業時 GPA*		N
	平均値	SD	平均値	SD	
2017	3.43	0.58	2.73	0.54	627
2018	3.53	0.53	2.70	0.61	157
2019	3.46	0.57	2.60	0.59	803
2020	3.37	0.67	2.75	0.55	155
2021	3.33	0.69	2.78	0.52	99

2022	3.20	0.74	2.71	0.61	105
------	------	------	------	------	-----

### 3.2 語と語の共起ネットワーク

GPA\*と年度で層別せず、データ全体（15回以上出現しかつ2.2で示した品詞計139語）について語と語の共起ネットワークを作成し、6年間で言及された主要な「満足度の理由」の把握を試みた。そのために、0.3以上の共起関係を示した語のみ描画した（図1）。その結果、「学生—生活」、「距離—近い」、「サークル—活動」「教養—科目」「単位—取得」の5つの共起パターンが抽出された。これらの語が意味するところ（例えば、「距離—近い」は、学修や研究指導における学生と教員との距離の近さを意味している）が、満足した主な理由であると考えられる。

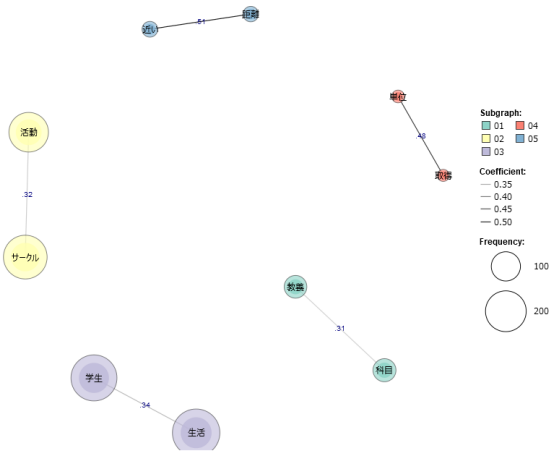


図1 相関係数0.3以上の語と語の共起ネットワーク

### 3.3 年度別の語の共起ネットワーク

出現する語を年度で層別し、語と年度の共起ネットワークを作成した（図2）。その結果、6年間の全ての年度で共通して出現していた語は「授業」であった。次点は「大学」と「学生」であった。したがって、コロナ禍によって学生生活に変化が生じて、卒業時の総合的満足度の主要な理由は「授業」および、「大学」と「学生」に関わることが示された。

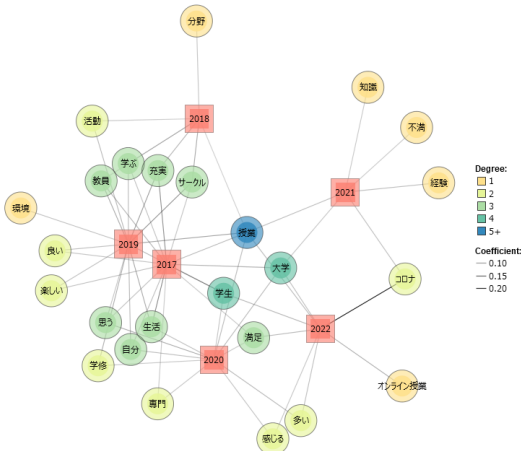


図2 年度別の語の共起関係  
一方、コロナ禍の影響が比較的大きいと思われる

2021年度と2022年度では、年度間の共通語に「コロナ」が出現し、特に2022年度は「コロナ」との関係が強くなっていた。また、2021年度と2022年度においては、他の年度と比較して、(1)出現する語が少ない、(2)他年度と共通して出現する語が少ない、といった特徴がみられた。

### 3.4 GPA\*別の対応分析

GPA\*の3グループと語の関係を対応分析にて示した（図3）。図3では、原点から遠い語ほど、グループにおける特徴的な語であることを示している。例えばGPA\*低群では、「単位」、「広い」、「過ごせる」が当該カテゴリにおける特徴的な上位3語であった。中群では「課外」、「取り組める」、「就職」、高群では「実習」、「参加」、「支援」「課題」であった。以上より、GPA\*低～中群では、「単位を取得する、友人と過ごす、課外活動をする、就活に取り組む」など、学修の周辺的な活動が満足度に関わる内容として記述される傾向にあった。一方、GPA\*高群では、「実習に参加する、学問・学修・専門などについて学ぶ・知る・相談する、教員と関わる」といった、学修の中心的活動が満足の理由にあげられていた。したがって、比較的高い満足度であっても、その理由はGPA\*の高低により異なる可能性が示唆された。

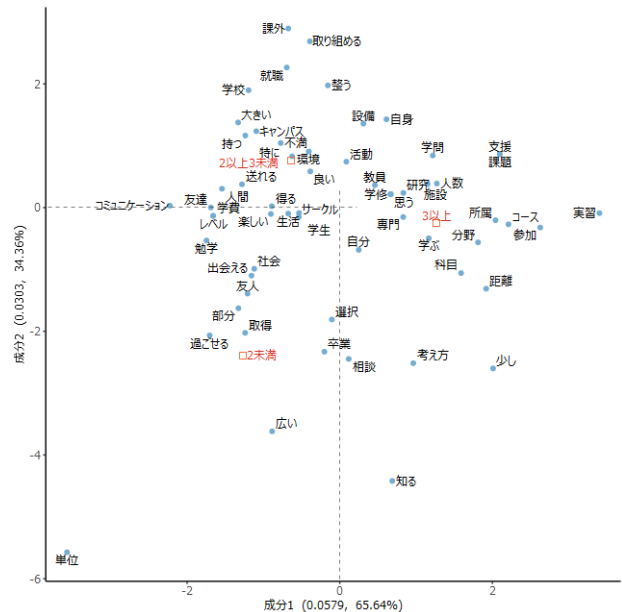


図3 GPA\*別の対応分析

### 付記

教学IRデータの利用および発表に際しては、教学IR委員会および研究倫理委員会の許可を得た。本研究は、科学研究費補助金基盤研究(B)（課題番号：24K00452）の支援を受けた。

### 参考文献

- (1) 樋口耕一: “KH Coder”, <https://khecoder.net> (参照2024.6.5)
- (2) 樋口耕一: “テキスト型データの計量的分析 —2つのアプローチの峻別と統合—”, 理論と方法, 第19巻, 第1号, pp.101-115 (2004)