

大学案内冊子の計量テキスト・階層分析による大学間特徴比較分析に関する研究

情報科学科 石原 弘将

指導教員：奥田 隆史

1 はじめに

近年、文部科学省では国立大学の機能別分化のために「世界的研究・教育の拠点」、「高度な教養と専門性を備えた人材の育成」、「職業実践能力の養成」の3類型を示した [1]。3 類型の中から自学の機能を選び、機能強化の方向性に沿った取り組み成果を運営費交付金の配分に反映する仕組みが制度化されている。同様の制度が公立・私立大学へも拡大されることを念頭に置いた議論がなされている。

このように、大学はそれぞれの機能別分化がなされようとしている。大学の機能や特徴が、具体的に大学案内冊子（以下、冊子）に含まれていれば、大学の広報を戦略的にこなうことができる。

一般的に、冊子には文章や写真を使った内容が含まれている。よって、大学が機能別分化しているか確認するためには冊子の文章や写真を分析し、大学の特徴が明確になっているか確かめる必要がある。

本研究では、冊子の計量テキスト分析と階層分析（AHP；Analytic Hierarchy Process） [2] による各大学の特徴を明らかにし、大学間特徴の比較分析をすることを目的とする。なお、紙面の都合上、階層分析による方法は卒業論文を参照していただきたい。

以下、第2節では、計量テキスト分析の概要、調査対象、分析手順を示す。第3節では、本学の経年変化、大学間特徴の差、大学間情報系学部特徴の差の分析結果を示す。第4節では本研究でのまとめと今後の課題を示す。

2 分析方法

計量テキスト分析：計量テキスト分析とは文章や名詞や動詞、形容詞などの品詞に分割し、それらの出現頻度や相関関係を分析することである [3]。

調査対象：本学と愛知県内の国立3校（A～C）、公立2校（D, E）、私立3校（F～H）の計9大学であり、各大学で作成された2019年度冊子を分析する。分析手順は以下のとおりである。

Step1: 各大学の冊子（pdf ファイル）を収集し、文章を抽出したテキストファイルにする。

Step2: テキストファイルを KHCoder [4] によって冊子の文章を単語ごとに分け、文章内での頻出回数に応じた頻出語リストを作成する。

Step3: 作成した頻出語リストを年度別や大学別などで比較し、分析をおこなう。

本研究での比較分析項目は（a）本学の経年変化、（b）2019年度の大学間特徴の差、（c）大学間情報系学部特徴の差である。なお、（a）は本学の2014、2016、2018、2019年度冊子の分析をした。

3 分析結果

(a) 本学の経年変化：各年度の上位10語の総単語頻出回数は2014年度が6786個、2016年度が1340個、2018年度が1545、2019年度が982である。これは2016年度以降の冊子には文章が減り、写真や画像が増えたからである。

2014年度の頻出語リスト上位10語（以下、上位10語）を基準とした頻出割合を図1に示す。なお、図1は各単語頻出回数を各年度の上位10語の総単語頻出回数で割り、正規化したものである。

図1より、各年度の上位10語は大きく変化しないことがわかった。

(b) 大学間特徴の差：本学とA～Hの上位10語を抽出し、頻出割合と種類を全大学間で比較した。比較結果より次のことがわかった。

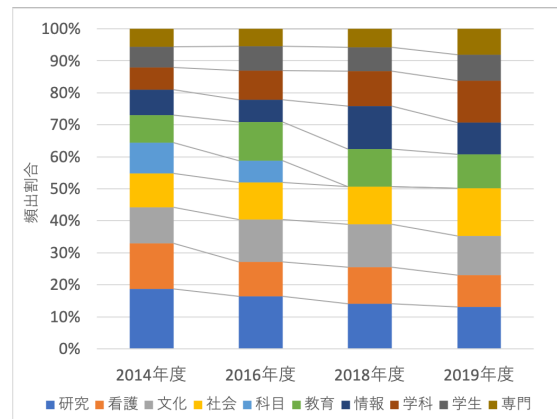


図1 本学の上位10語の頻出割合経年変化

- (1)：全大学で「研究」のみが共通して使用されている。「研究」は6つの大学で最も多く使用されている。
- (2)：全大学で上位10語の種類が異なっている。例えば、Bでは「工学」、「技術」、「材料」などの工学分野、Dでは「音楽」、「美術」、「芸術」などの芸術分野、Fでは「文化」、「留学」、「国際」などの語学分野に関する単語が多く使用されている。
- (3)：国立、公立、私立を明示的に区別できる単語の使用はなく、上位10語だけでは国立、公立、私立の特徴判別はできない。

(c) 大学間情報系学部特徴の差：本学とA, B, G, Hの上位10語を抽出し、本学を基準として頻出割合と種類を全大学間で比較した。比較結果より次のことがわかった。

- (1)：全大学で「情報」、「技術」、「システム」、「研究」の4語が共通して使用されている。
- (2)：AとBで使用されている単語の種類は全く同じである。同様にGとHで使用されている単語の種類はほとんど同じである。
- (3)：AとBにおける頻出割合が多い単語は、GとHでは少ない。同様にGとHにおける頻出割合が多い単語は、AとBでは少ない。

4 おわりに

本研究では、冊子の計量テキスト分析による本学の経年変化と大学間特徴を比較分析した。その結果、本学の経年変化では使用される単語はほとんど変わらず、総単語頻出回数が減ることがわかった。大学間特徴の差では各大学の特徴に関する単語が多く使用されており、全調査大学の特徴が明確であることがわかった。大学間情報系学部特徴の差では国立と私立で使用される単語の種類と頻出割合が違うことがわかった。

今後の課題は愛知県内の大学だけでなく他の地域においても同様な結果が得られることができるかなどがある。

参考文献

- [1] 文部科学省，“第2章 新時代における高等教育の全体像：文部科学省”，http://www.mext.go.jp/b_menu/shingi/chukyo/chukyo0/toushin/attach/1335594.htm，閲覧日：2019年1月17日。
- [2] 石原他，“AHPによる大学案内冊子掲載写真の決定手法の提案”，日本オペレーションズ・リサーチ学会2018年秋季研究発表会，pp.230-231，2-F-1，2018。
- [3] 樋口耕一，『社会調査のための計量テキスト分析—内容分析の継承と発展を目指して』，ナカニシヤ出版，2014。
- [4] KHCoder，“KHCoderの主な機能と分析手順”，<http://khcoder.net/diagram.html>，閲覧日：2018年12月14日。