

論文 / 著書情報
Article / Book Information

論題(和文)	建設系大学院生を対象としたWeb会議システムを用いた地震時における安心についてのワークショップ その2 テキストマイニングを用いた共起ネットワーク分析
Title(English)	Online workshop on relief at the time of earthquake for graduate students majoring architecture and civil engineering: (Part 2) Co-occurrence network analyses using a text mining method
著者(和文)	鈴木一徳, 佐藤大樹, 永岑光恵
Authors(English)	Kazunori Suzuki, Daiki Sato, Mitsue Nagamine
出典 / Citation	日本建築学会関東支部研究報告集, II, , pp. 469-472
Citation(English)	, II, , pp. 469-472
発行日 / Pub. date	2021, 3

建設系大学院生を対象とした Web 会議システムを用いた

地震時における安心についてのワークショップ

その2 テキストマイニングを用いた共起ネットワーク分析

教育—専門教育

正会員 ○ 鈴木一徳^{*1}正会員 佐藤大樹^{*2}〃 永岑光恵^{*3}

安心 安全 耐震 建設系大学院生 ワークショップ 共起ネットワーク

1. はじめに

本報その1では、グループワーク（GW）の概要について報告した。本報その2では、GW（本報その1のGW②とGW④）で得られた回答をKH Coder⁶⁾というテキストマイニングに特化したソフトウェアを用いて共起ネットワーク分析を行った結果を報告する⁴⁾。分析対象のデータは、GW②「高層建築における安心とは（本報その1の4.2節）」およびGW④「何が地震時の建物内で不安を喚起させるのか（本報その1の4.3節）」で提出された成果物である。

2. テキストマイニング：KH Coder

2.1 KH Coder の概要

KH Coder は、文章データを統計的に分析する計量的テキスト分析を行うために開発・公開されたフリーのソフトウェアである⁹⁾。KH Coder を用いた研究は多岐に渡っている。例えば、SNS 上のコミュニケーションの研究、新聞や雑誌記事のテキスト分析、国会や裁判などの会議録のテキスト分析を基にした意思決定プロセスの研究、自由記述アンケートのテキスト分析などがある。KH Coder を用いて行われたテキストマイニングにかかわる研究はすでに1500件を超えているという⁷⁾。

本来、KH Coder は、小説や自由記述アンケートなどの長文のテキストを分析するために用いられるものであるが、本研究ではGW成果物として提出された短文回答への適用を試みた（具体的な成果物は本報その1の図5および図6を参照のこと）。テキストマイニングにより文章データを可視化することで、主観を排除し、客観的にデータを分析することができる。

2.2 KH Coder の設定と分析手順

GWの成果物は、テキスト化を行った。元データの作成に際しては、1つの回答（キーワード・短文）を1行に入力した。ただし、全部で15グループあったが、1つのグループについては、ホワイトボードのデータの保存が

できていなかった。そのため、本報その2では、保存できなかった1つのグループ以外の14のグループのデータを分析の対象とした。

GW②「高層建築における安心とは」では、D系から117の回答、E系から82の回答、合計で199の回答が得られた。また、GW④「何が地震時の建物内で不安を喚起させるのか」では、D系から125の回答、E系から81の回答、合計で206の回答が得られた。どちらのGWでもE系よりもD系の方が回答数が多いが、これは本報その1で報告した通り、D系の方がE系よりも学生数が多いことが影響している可能性がある。D系・E系の1人当たりの平均回答数を算出したところ、それぞれ6.9（D系）、6.3（E系）であったことから、D系の方が1人当たりの平均回答数がやや多いことが窺える。

データ分析に先立ち、テキストデータを形態素解析した。形態素解析をすることで、共起ネットワーク分析を行う際に、内容語（名詞、動詞、形容詞）と機能語（助詞）を区別し、内容語のみで共起ネットワークを作成することができる。

共起ネットワーク分析を行う際に、D系およびE系それぞれの回答の傾向（共通点・相違点）を可視化するために、各回答にD系またはE系のインデックスも外部変数として読み込んだ。

2.3 共起ネットワーク分析

共起ネットワーク分析を行うことで、語と語の結びつき（共起関係）を可視化することができる。

次節以降の図中の円は抽出された語を表している。円の大きさが語の出現頻度を示しており、出現頻度が高くなると円も大きくなる。本報その2では、比較的強く結びついている語同士を自動的に検出してグループ分けを行うサブグラフ検出を使用した。図中の線は共起関係の強さを表している。線の濃さが共起関係の強さを示しており、強い結びつきであれば線が濃くなる。また、共起

関係を表す線には、標準化された係数（Jaccard 係数）が記されている。Jaccard 係数は、係数は特徴の程度を示しており、0.0～1.0 の間で示され、0.1 以上で弱い特徴、0.3 以上で強い特徴があるとされている。

図中の四角で囲まれた「Design」と「Engineering」は、外部変数を表している。抽出された語が D 系と E 系に共通するものであるのか、または D 系・E 系それぞれに特徴的なものであるのかを示している。

なお、本報その 2 では、2020 年 12 月 17 日にリリースされたバージョンの KH Coder (3.Beta.02c) を用いてデータ分析を行う。

3. 分析結果：GW②「高層建築における安心とは」

3.1 全体の結果

GW②で得られた 199 の回答を形態素解析した結果、総抽出語数 1432 語、異なり語数 407 語が抽出された。図 1 は、D 系・E 系両方の回答について、最小出現数を「3」で共起ネットワークを表す図を作成したものである。

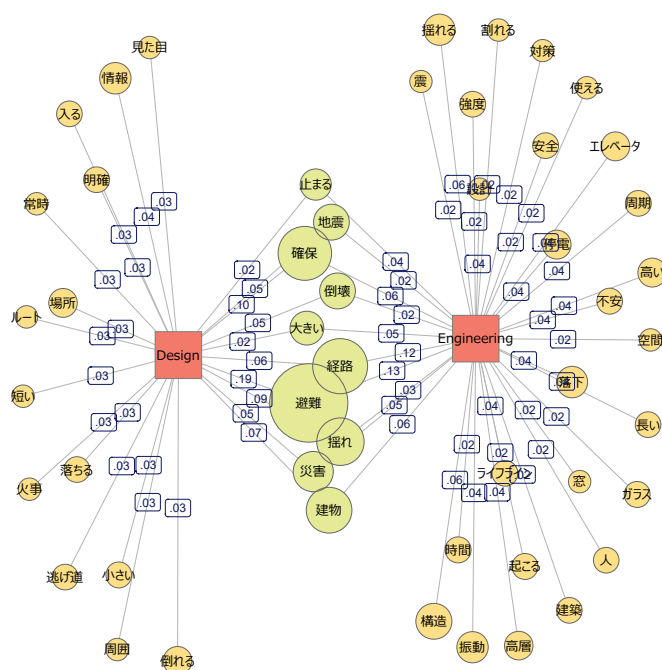


図 1 GW②の D 系・E 系の共起ネットワーク

D 系・E 系に共通して得られた回答は、「止まる」「地震」「確保」「倒壊」「大きい」「経路」「避難」「揺れ」「災害」「建物」の計 10 の語であった。また、D 系に特徴的な回答の例としては、「情報」「場所」「落ちる」「逃げ道」「小さい」「倒れる」など、計 14 の語が抽出された。さらに、E 系に特徴的な回答の例としては、「揺れる」「エレベーター」「停電」「高い」「落下」「振動」「構造」「建築」など、計 26 の語が抽出された。

図 1 より、「避難」が D 系・E 系の両者に共通し、且つ

出現頻度も最も多いことが分かる。つまり、「高層建築における安心」が「避難」であることが示されている。この結果は、本研究プロジェクトの前回の WS 実施時とも一致する結果である⁴⁾。また、D 系に特徴的な語として「情報」が抽出されたこと、E 系に特徴的な語として「建築」が抽出されたことは前回の WS の結果とも一致する⁴⁾。

3.2 D 系・E 系それぞれの結果

D 系・E 系それぞれの共起関係を可視化し、専門性による視点の相違点を見ていく。

GW②で得られた D 系の 117 の回答を形態素解析した結果、総抽出語数 859 語、異なり語数 285 語が抽出された。図 2 は、D 系の回答について、最小出現数を「2」で共起ネットワークを表す図を作成したものである。また、GW②で得られた E 系の 82 の回答を形態素解析した結果、総抽出語数 573 語、異なり語数 236 語が抽出された。図 3 は、E 系の回答について、最小出現数を「2」で共起ネットワークを表す図を作成したものである。

図 2 と図 3 を比べると、D 系にあるが E 系にはない語の共起ネットワークとして、「ライフライン」「切れる」、「飲料」「予備」、「情報」「不足」「入る」「見える」、「不安」「周り」「様子」、「建物」「知る」、「ガラス」「割れる」、「煙」「出る」、「エレベーター」「閉じ込める」が挙げられる。他方、E 系にあるが D 系にない語の共起ネットワークとして、「物」「家具」「倒れる」、「揺れる」「時間」、「恐怖」「感じる」「死ぬ」が挙げられる。

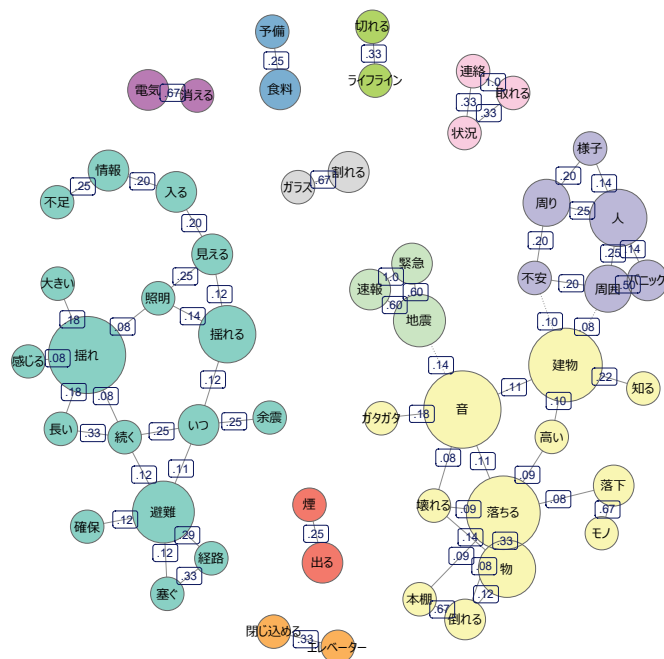


図 2 GW②の D 系の共起ネットワーク

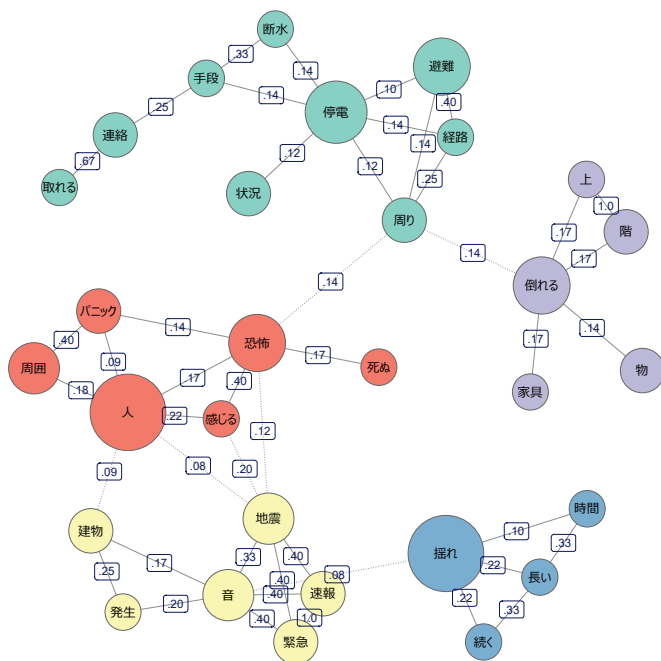


図3 GW②のE系の共起ネットワーク

4. 分析結果：GW④「何が地震時の建物内で不安を喚起させるのか」

4.1 全体の結果

GW④で得られた 206 の回答を形態素解析した結果、総抽出語数 1153 語、異なり語数 347 語が抽出された。図 4 は、D 系・E 系両方の回答について、最小出現数を「3」で共起ネットワークを表す図を作成したものである。

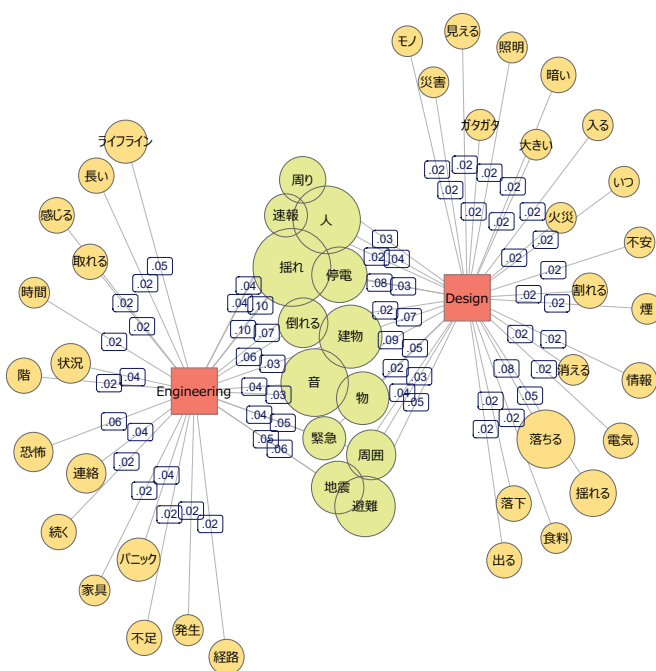


図4 GW④のD系・E系の共起ネットワーク

D系・E系に共通して得られた回答は、「周り」「速報」「人」「揺れ」「停電」「倒れる」「建物」「音」「物」「緊急」「周囲」「地震」「避難」の計13の語であった。また、D系に特有の回答の例としては、「揺れる」「落ちる」「情報」「食料」「見える」など、計21の語が抽出された。さらに、E系に特有の回答の例としては、「ライフライン」「パニック」「恐怖」「連絡」など、計15の語が抽出された。

図4より、「揺れ」「音」「停電」「人」などがD系・E系の両者に共通し、且つ出現頻度も多いことが分かる。つまり、「地震時の建物内で不安を喚起させるもの」が「揺れ」「音」「停電」「人」などであることが示されている。

4.2 D系・E系それぞれの結果

D系・E系それぞれの共起関係を可視化し、専門性による視点の相違点を見ていく。

GW④で得られた D 系の 125 の回答を形態素解析した結果、総抽出語数 631 語、異なり語数 234 語が抽出された。図 5 は、D 系の回答について、最小出現数を「2」で共起ネットワークを表す図を作成したものである。また、GW④で得られた E 系の 81 の回答を形態素解析した結果、総抽出語数 522 語、異なり語数 226 語が抽出された。図 6 は、E 系の回答について、最小出現数を「2」で共起ネットワークを表す図を作成したものである。

図5と図6を比べると、D系にあるがE系にはない語として、「訓練」「情報」「伝達」、「被害」「倒れる」「少ない」、「揺れ」「小さい」が挙げられる。他方、E系にあるがD系にはない語として、「落下」「防止」、「安全」「空間」「設計」、「建築」「部材」が挙げられる。

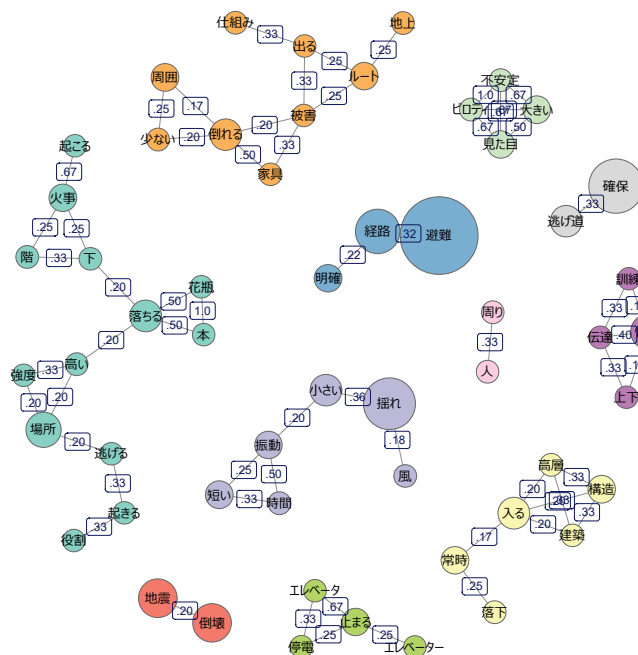


図5 GW④のD系の共起ネットワーク

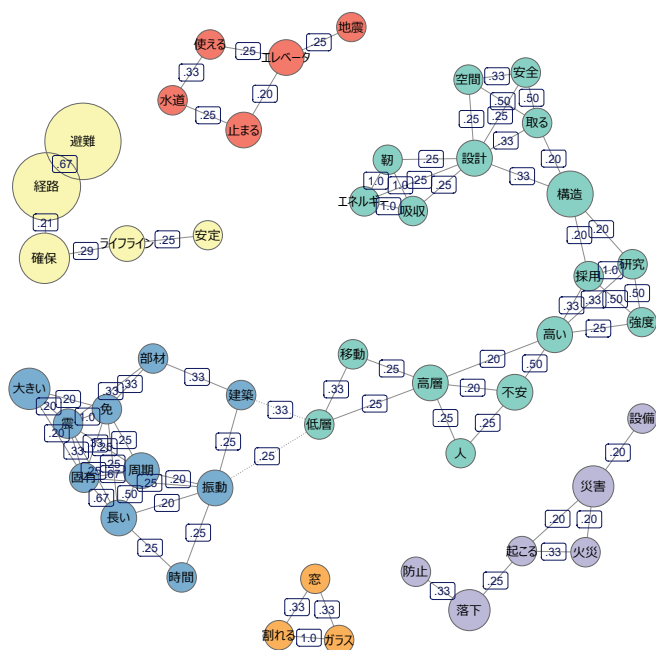


図6 GW④のE系の共起ネットワーク

5. まとめ

本報その2では、建設系大学院生を対象に行ったWSでの成果物を対象に、KH Coderを用いて共起ネットワーク分析を行い、D系とE系の共通点・相違点を探ることを目的としてきた。データ分析の結果、以下の5点が明らかになった。

- (1) 対面授業で実施したWSとWeb会議システムで実施したWSの結果の比較から、WSの実施方法に拘わらず同様の結果が得られた(3.1節)
- (2) GW②に関して、D系とE系に共通する語の代表は「避難」であり、両者にとって「安全＝避難」という認識であった(3.1節)
- (3) GW②に関して、D系に特徴的な語のネットワークは「ライフライン」「切れる」「飲料」「予備」「情報」「不足」「入る」「見える」など、ソフト面に関する語が抽出されたが、E系に特徴的な語のネットワークは「物」「家具」「倒れる」「揺れ」「時間」など、ハード面に関する語が抽出された(3.2節)
- (4) GW④に関して、D系とE系に共通する語の代表は「揺れ」「停電」「建物」「倒れる」「音」であった(4.1節)
- (5) GW④に関して、D系に特徴的な語のネットワークは「訓練」「情報」「伝達」など、ソフト面に関する

る語が抽出されたが、E系に特徴的な語のネットワークは「落下」「防止」「建築」「部材」など、ハード面に関する語が抽出された(4.2節)

以上、2種類のGWに関して、建設系大学院生の専門性(D系・E系)の認識の共通点および相違点を、共起ネットワークという分析手法を用いて概観し、おおむね各々の専門性に応じた特徴的な語を炙り出した。しかし、テキストマイニングという計量的な手法を用いて、信頼性の高いデータ分析を行うには、今回対象としたデータでは、データの大きさという観点からは十分であるとは言えない。今後は、GWを経て行った自由記述課題の分析を行うことで、D系・E系の特徴を見極めることができる可能性がある。

謝辞

本研究の一部は、JST 産学共創プラットフォーム共同研究推進プログラムによるものです。また本ワークショップを実施するにあたり、東京工業大学博士課程の南健斗氏、藤田悠氏に多大なご協力をいただきました。ここに記して謝意を表します。

参考文献

- 1) 東京工業大学 OPERA 社会活動継続技術コンソーシアム (SOFTech): <http://www.softech.titech.ac.jp/> (2019.6 参照)
- 2) 佐藤大樹, 永岑光恵, 鈴木一徳: 建設系大学院生を対象とした地震時における高層建築の安全・安心に関するワークショップ その1: 専門性に関するアンケート結果, 日本建築学会大会学術講演梗概集, 教育, 13011, pp.21-22, 2020.9
- 3) 永岑光恵, 佐藤大樹, 鈴木一徳: 建設系大学院生を対象とした地震時における高層建築の安全・安心に関するワークショップ その2: グループワークの全容, 日本建築学会大会学術講演梗概集, 教育, 13012, pp.23-24, 2020.9
- 4) 鈴木一徳, 佐藤大樹, 永岑光恵: 建設系大学院生を対象とした地震時における高層建築の安全・安心に関するワークショップ その3: テキストマイニングによる計量的分析, 日本建築学会大会学術講演梗概集, 教育, 13013, pp.25-26, 2020.9
- 5) 川喜田二郎: 発想法 創造性開発のために, 中公新書, 1967
- 6) 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
- 7) 樋口耕一: 計量テキスト分析およびKH Coderの利用状況と展望, 社会学評論, 68(3), 2017

*1 東京工業大学 環境・社会理工学院 博士後期課程 修士(文学)
 *2 東京工業大学 未来産業技術研究所 准教授・博士(工学)
 *3 東京工業大学 リベラルアーツ研究教育院 准教授・博士(理学)

*1 Graduate student, School of Environment and Society, Tokyo Institute of Technology, M. A.
 *2 Assoc. Prof., FIRST, Tokyo Institute of Technology, Dr. Eng.
 *3 Assoc. Prof., ILA, Tokyo Institute of Technology, Dr. Sc.