

IV-29

土木用語の構造化に関する基礎的研究

北海道大学大学院 学生員 森 若 峰 存

北海道大学工学部 正 員 佐 藤 馨 一

北海道大学工学部 正 員 五十嵐日出夫

1. 本研究の目的・特徴

近年、土木工学のフィールドは、人の日常生活において直接、目の触れることのできる都市内が主流になりつつある。さらに、地球環境・自然保護の声の高まりなど、社会の変化は豊かな暮らし、存続可能な経済を求める方向に向かっている。そして、我々の土木工学は「開かれた土木」になることを要求されている。

しかし、物事を理解するうえで、最も重要であるものが言葉であるにもかかわらず、我々の言葉である「土木用語」は理解しにくいものである。そのためもあって、土木技術者と住民との意志疎通の障壁が非常に高い。さらに、同じ土木工学でありながら、例えば河川工学の専門家が、道路工学で用いられている専門用語を理解できない。このように土木用語は限られた一部の専門家でなければ「本当の」意味で理解されていないのが実状ある。

このように、広く住民に土木工学を理解していただくためには、言葉が重要であり、本研究は土木用語の構造化を行うことを目的とする。また、最終的な図を「リレーションチャート (Relation Chart)」（図-6）と名付け作成する。

本研究におけるリレーションチャートを作成するためのフローチャートを図-1に示す。本研究では、あらかじめ土木用語がいくつか抽出されていることを前提して、図中の用語の抽出については触れていない。この流れの中で特に力を置いた点が、用語の関係付けの方法である。これを、簡単に述べると、2つの用語の関係が「ある／ない」を判断するにあたり、その用語を取り込んだ文が成立「する／しない」で判断している。

「文が成立する」	→	「関係あり」
「文が成立しない」	→	「関係なし」

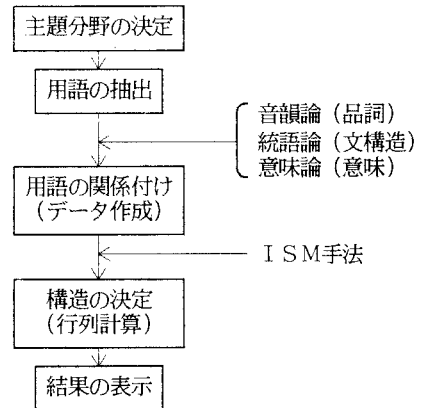


図-1 リレーションチャート作成のためのフローチャート

ここで、基本となる文型を設定するために、言語学の考え方をいくつか利用している。

また、もう1つの柱としてI S M手法の適用がある。

土木用語の構造は複雑なものであるため、本研究ではI S M手法を用いて、構造を決定する手助けにしている。そのために、行列計算をするためのデータの作成の作業において工夫をする必要があり、関係が「ある／ない」を判断するために、「文」を用いた。

2. 土木用語を構造化することの意義

ある言葉があった場合、理解できない理由を分析すると次項の図-2のようになる。

②は本来問題にできないことであるので、ここでは取り上げる必要はないので、①について考える。

家庭において日常用いられている言葉が理解できない理由は主に単語レベルの問題である。例えば、1つの単語で意味が複数ある場合（あいまいさ）、定義が明確にされていない場合（不確かさ）である。しかし、専門用語は、構造レベルの理由が非常に大きい。

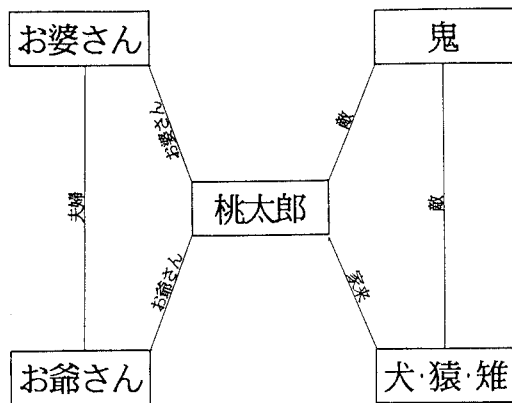
A Study on the Structural Modeling of Civilengineering Terms

by Mineari MORIWAKA, Keiichi SATHO and Hideo IGARASHI

- ①言葉自体に問題がある場合
 - 定義が不明確（単語レベル）
 - あいまいさ
 - 不確かさ
 - 定義が複雑（構造レベル）
 - 定義の中に知っている単語が多い
 - 定義の中に知らない単語が多い
- ②人に問題がある場合
 - 理解する必要が無い（無関心）
 - 理解する努力が足りない（怠惰）
 - 理解する能力が無い

図－2 言葉が理解できない理由

構造レベルの話をするうえで、物語の『桃太郎』を例として考える。登場人物としては図－3のように「桃太郎」「鬼」などいくつかの単語が登場する。いくなれば、これらの登場人物は『桃太郎』という専門分野で用いられている専門用語と考えることができる。ただ、この場合は物語が理解しやすく、それぞれの用語の間の関係が、図－3のように日本人ならば誰でも頭の中で関係づけの整理ができています。つまり、単語レベルと構造レベルの両レベルで理解されやすい。



図－3 『桃太郎』中の登場人物の関係

しかし、土木用語は両レベルとも理解しにくい言葉である。なぜならば、土木工学は極めて専門的であるために単語を並べられても整理ができない。つまり、物語のように、構造レベルの問題が非常に大きい。

ここで、専門用語について考える。「用語（専門用語）」とは、「使用する言語・語句。特定の部門や人に特に使われることば。」である。つまり、ある用語とはその分野に携わる者（専門家）であれば知っているが、専門外の者にとっては、容易には理解しがたい

言葉である。このことを考慮すると、ある専門用語を理解しようとした時、最も厄介な点は、用語と用語の係わりを理解しなければならないことである。また、日常の用いられる言葉と違う点でもある。日常の言葉は1つの単語をそのまま国語辞典に載っているように理解できる。一方、専門用語というものは、1つの単語を理解したからそれで良いというものではない。例えば、土質力学において、「静止土圧」という単語のみを理解しようとすると、不十分である。「静止土圧」と対になっている「受動土圧」を併せて理解しなければ「本当の意味」で理解したことにならない。このように土木用語に限らず専門用語というものは、個々の専門用語を1つ1つ理解するだけでは不十分であり、いくつかの専門用語との関連を考えることで、用語を理解できる。つまり、用語と用語のつながりを明確にし構造化することは、土木工学を理解するうえで有効である。

3. シソーラス

言葉を構造化する既存の方法としてシソーラスがある。シソーラスとは「（知識の宝庫としての）辞典、百科辞典、宝典；特に同意語反異語辞典：宝庫：（コンピュータに関する）シソーラス、類語辞典、コンピュータに記憶されている情報に関する索引；キーワードを与えることによって必要な情報を取り出すことができるよう、題目の包括的な目録からなっている。」と定義されている。つまり、広義の意味では、言葉（知識）が集まっているものを指し、狭義では類語辞典となる。このような中で、シソーラス固有の特徴は、収録した用語間の構造的な関係を識別し表示することにある。

シソーラス構築の手順としては、大きく次のようになっている。

1. 主題分野を決定する。
2. 用語を抽出する。
- (3. 用語をいくつかの基本カテゴリーに分類する)
4. 各基本カテゴリーの中で用語の関係付けをする
5. 結果を表示する。

上のように、主題分野を決定し、その主題分野に關する用語を抽出し、いくつかの基本カテゴリーに分類し、各基本カテゴリー中の用語の関係付けを行い構造化している。

この関係付けの作業で、シソーラスは、階層関係に注目している。階層関係とは、「上位および下位のレ

ベルを示す。上位語は類あるいは全体を表し、下位語はその1要素あるいは1部分を表すものである。つまり、階層関係は、上位概念および下位概念を基本として関係付けられる関係である。

また、シソーラスと構造化されていない用語リストとを区別する基本的な特徴であり、再現性を高め適合性をも高める重要な要素である。

実際の例として、図-4のように、主題分野を「セメント」に設定し、基本カテゴリーを1つ（セメント）に決定した場合のものを示す。これを見て分かるように、上下関係が非常にはっきりしており、「セメント」の下位概念として、「ポルトランドセメント」「混合セメント」などがあり、さらに下の下位概念として、「早強ポルトランドセメント」などがあ

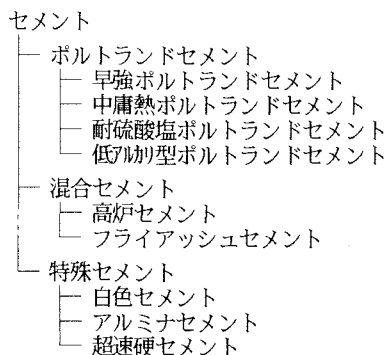


図-4 シソーラスの例（「セメント」）

このようにシソーラスでは、階層関係に注目しているわけであるが、そのため上下関係がはっきりしているものについてはたいへん有効な方法である。しかし、本研究のリレーションチャートの例として図-6の「土地区画整理事業」のように、上下関係では説明できない用語間の関係付けにおいては、上位概念や下位概念がはっきりしていないために、シソーラスによる構造化は不可能である。

4. リレーションチャートの作成

4-1. 言語学の体系

本研究では用語と用語の関係付けを行う際に、図-1のフローチャートのように言語学の考え方をいくつか利用しているのであるが、この節では言語学について概要を説明する。言語学における研究分野は次のように大きく3つある。

①音声・音韻		音声論・音韻論
②文法		形態論・統語論
③意味		意味論

音声学（phonetics）では、言語音を作り出すときには、どの器官がどのような運動を行っているのか。また、そのようにしてできあがった言語音が、ひとたび空気中に放出された後は、どのような音響構造としてこれが捉えられるか。さらには、相手の鼓膜を刺激した音波は、どのようなメカニズムによって大脳で意味単位として知覚されるのかを研究する分野である。

音韻論（phonology）では、音声学による観測を通して得られた知見をもとにして、言語音全体を律する機能的側面を構造と体系の両面から推論する分野。

形態論（morphology）では、あるまとまった社会慣習の意味と対応している最小の言語単位を形態素（morpheme）と呼び、形態素の抽出とその諸特徴を他の分野の文法的諸単位から独立させて論じる分野。

統語論（syntax）では、1つの言語における的確な文を作り出し、同時に不的確な語の羅列を排除する仕組みを明らかにすることを目的とした分野。

意味論では、文字どおり意味についての分野である。

この中で、本研究では音韻論から品詞について、統語論から日本語の文構造について、意味論から意味について考え方を利用している（図-1）。

4-2. 動詞について

用語と用語の関係が「ある／ない」を論じるにあたって、品詞の中の動詞に特に注目している。

言語学の音韻論（品詞）の分野から、日本語の動詞を意味によって分類すると、存在をあらわす「ある」、行為や作用をあらわす「する」（行為動詞）、過程や状態の変化をあらわす「なる」（自発動詞または自発動詞）の3つに大別される。

また、日本語の行為動詞と自発動詞は対をなしているものが多い。「こぼす／こぼれる」、「うつす／うつる」、「たすける／たすかる」などが対の例である。このように、自発動詞の中には、行為動詞にある程度、集約できるものがある。文では、「動作主＋被動者＋行為動詞」の構造になっている。行為動詞ではなく自発動詞を用いると、「被動者＋動作主＋自発動詞」の構造になっている。しかし、主語は変わるが動作主は変わらない。例として、「一郎は次郎に数学を教える。」「次郎は一郎に数学を教わる。」と同じ意味で2つの

文章を作ることができる。この2つの違いは「は／に」である。それとともなって前者は行為動詞（教える）、後者は自発動詞（教わる）となっている。そこで、本研究では、自発動詞の中で行為動詞と成る動詞は、行為動詞の形で表して用いることを原則とする。つまり、文は能動態を採用し、この例では前者を採用する。

また、「ある（存在動詞）」は、例えば「花がある。」という文型になって、用語間の関係を議論することを使うことはできないので、除外する。

さらに、日本語の動詞には形式動詞がある。例えば、「である」「です」「であります」「でございます」「だ」などがある。これらはすべて語源が断定の意味を持っている「である」であり、これが用語と用語の関係を定義するうえで重要である。

まとめると、用語と用語との関係の仲立ちとなる動詞は「する（行為動詞）」と「なる（自発動詞）」、「である（形式動詞）」である。

4-3. 日本語の構造

日本語の構造は、言語学における統語論において研究が進められている。基本となる構造は、次のようになっている。

「名詞」が「名詞」を「動詞」。

この文が基本となって、日常行われている会話文や小説の文などの複雑な文が出来上がっている。

4-4.. 基本文型の設定

先ほどの「する（行為動詞）」「なる（自発動詞）」と断定の形式動詞「である」の3つを組み合わせ、前述の日本語の文構造を考慮し、次の3つの基本文型を設定した。

①「AはBの……である。」
形式動詞

②「AはBを（に）……する（するための）……である。」
行為動詞 形式動詞

③「AはBになる……である。」
行為動詞 形式動詞

さらに、例えば「減歩」のような抽象名詞は後に「する」を付け、「減歩する」のように用語自体が動詞になるものがある。そこで、②は次のように②'に変化する。

②'「AはBする（するための）……である。」
行為動詞 形式動詞

5. 実際の土木用語の構造化例

5-1. I S M手法適用のためのデータ作成

I S M (Interpretive Structural Modeling) 手法が開発された目的は、①複雑なシステムを構成している多数は要素間に存在する何らかの因果関係をコンピュータとの対話によって解明していくこと、②グループ意志決定の場において、当初統一したい件がない場合に、I S M手法を介して除々に合意を形成してゆき、グループの政策決定に役立てることなどである。

このI S M手法を土木用語の構造化の計算過程で用いることは有効である。しかし、従来のI S M手法においては、行列計算をさせるためのデータを、なんとなく関係が「ある／ない」によって作成していた。一方、本研究では言葉を構造化するという従来のI S M手法の利用方法と違うため、基本文型を用いて、関係の「ある／ない」を判断している。

5-2. 「土地区画整理事業」についての構造化

ここでは例として「土地区画整理事業」について考える。まず、「土地区画整理事業」に関する周辺の土木用語をいくつか取り上げる。そして、用語と用語の関係付けを1つ1つ行うわけであるが、実際の関係付けは図-5のフローチャートに従って行う。

さて、図の上段から中段にかけての作業では、実際の作業として、ある用語と用語の関係が「ある／ない」を基本文型に入れることをするのであるが、省力化をするために、用語を下の3つに分けた。

- ①人名
- ②人名以外の形のある用語（具象名詞）
- ③人名以外の形のない用語（抽象名詞）

このように分けることによって、用語と用語の組み合わせによっては適用可能な基本文型が減る場合がある（中段）。さらに、抽象名詞は、例えば「減歩」のように、「する」を語尾に接続すると「減歩する」というサ変動詞に変化することも考慮している。

説明として、「減歩」と「保留地」についての関係付けを判断することを取り上げる。まず、2つ用語を上に従って分類すると、「減歩」は抽象名詞であり、「保留地」は具象名詞である。これは、フローチャートの上段で処理している。中段で、適用可能な基本文型は①、②、②'があることが決定する。結果として②を採用し、以下のような文を作った。行為動詞

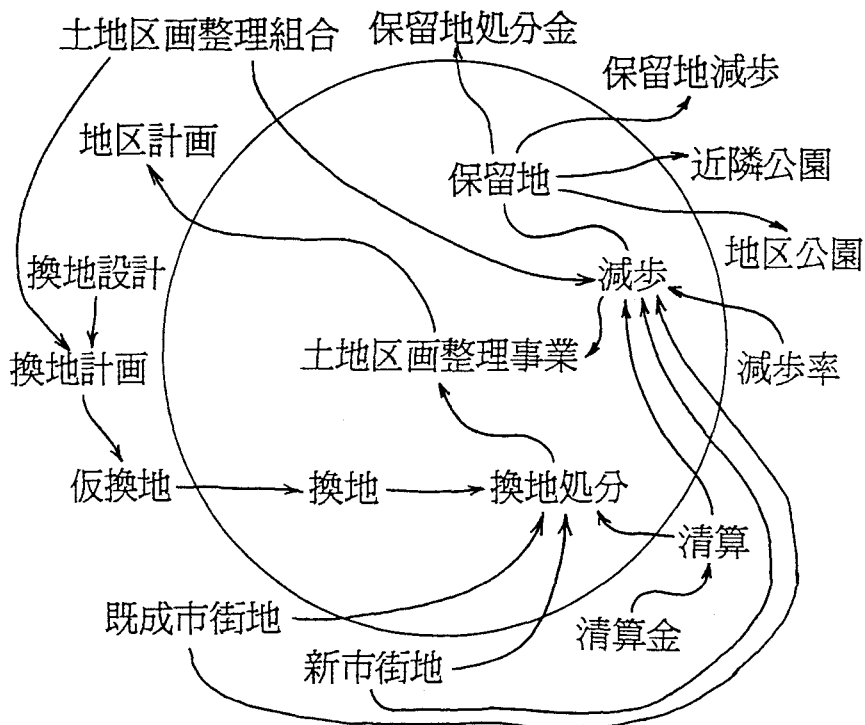


図-6 「土地区画整理事業」についてのリレーションチャート

(「作る」)を介在して文が成立している。

減歩は保留地を作るための方法である。
A B 行為動詞 形式動詞

最後に、下段で次のように矢印を引く。

「A」 → 「B」
「減歩」 → 「保留地」

このように、19項目について、全ての組み合わせについて図-5のフローチャートに従ってデータを作成し、ISM手法の行列計算を行った。

アウトプットとして、19項目のうち重要と思われるいくつかを中心において、図-6のように最終的なリレーションチャートを作成した。

6. まとめ

図-6のように、「土地区画整理事業」に関する土木用語の構造化を行うことにより、用語と用語の関係が整理された。そして、土木用語のように、きわめて専門的で理解しにくい専門用語を構造化するためには、そもそも「言語を構造化する」という困難な作業を基本文型を設定し、「文」によって関係を明らかにする

ことで可能になった。また、構造を決定するうえで用いたISM手法の適用が可能になった。また、なぜ関係があるとしたかが「文」の形で残るため、なぜ関係があるとしたかが、いつ・誰にでも理解できる。今後の課題として次のような点が重要になると考えられる。

- ・図-1の単語レベルで問題を抱えている用語の扱い方を考慮すること。
- ・漏れのない用語の抽出の方法の確立すること。
- ・基本文型に入れて、文が成り立つかどうかを判断する人を選出すること。
- ・用語の組み合わせ数が多いため関係付けの回数が多いこと(図-6の場合、19項目なので171通り)。
- ・抽出された用語の中で重要なものを選び出すこと。
- ・結果の表示において用語の配置により、見栄えや読み手が受ける感じが変わってしまうこと。

参考文献

- 1) Jean Aitchison; Alan Gilchrist 著; 内藤衛亮; 中倉良夫; 影浦映ほか訳: シソーラス構築法, 丸善(1989)
- 2) 中島文雄: 日本語の構造, 岩波新書(1987)
- 3) 石黒昭博: 現代言語学序説, 山口書店(1984)