

漸近的对応語彙推定法を用いた 異言語テキスト間の概念ネットワーク

村井源

東京工業大学 大学院社会理工学研究科

複数の言語で書かれた古典テキスト群において、異言語間の概念の関係性をより明示的に把握するため、古典テキストとその現代語の翻訳に対して漸近的对応語彙推定法を適用し、対応する語の組を抽出し、関係性を可視化する手法を提案した。聖書のヘブライ語とギリシア語原典に対して提案手法によるケーススタディ的分析を行い、各翻訳の共通性や個別的解釈の特徴を計量的に抽出しネットワーク構造を用いて視覚化した。今後の課題としては、得られた語の対応関係性の数値性を活かした数値的意味解析や構造化手法の確立や、分析結果のデジタルアーカイブとの統合が考えられる。

Semantic Networks between Texts in Different Language Based on the Asymptotic Correspondence Vocabulary Presumption Method

Hajime Murai

Graduate School of Decision Science and Technology
Tokyo Institute of Technology

In order to clarify the relationship between the concepts of different languages in classic texts those written in multiple languages, a method was proposed to visualize those relationships. By utilizing the Asymptotic Correspondence Vocabulary Presumption Method to classic texts and those modern translations, the proposed method enables to extract corresponding vocabulary pairs and to visualize by making a network. The Bible was analyzed as a case study for the proposed method. As a result, common and individual characteristics of modern translations' interpretation were extracted and visualised in forms of network. As a future works, quantitative characteristics of extracted relationships should be utilized for numerical semantic analysis and structural analysis. Integration of digital archive system and results of the analysis is also another future work.

1. はじめに

古典文献の研究において翻訳の問題は様々な形で関わってくる。当然のことながら多くの読者は古典語に習熟しておらず最初の段階では古典文献に翻訳を通して触れることが一般的である。また古代世界においてもグローバルな知識の交流が盛んであり、研究対象となる特定のテキストは、それが書かれた言語以外の多数の言語のテキストを引用・参照して記されていることが少なくない。このような他言語テキストからの引用・参照においては何らかの形で翻訳が介在しており、テキスト解釈を深めるうえでは翻訳の利用は避けて通ることができない。

しかしながら翻訳において別言語に元言語の意味を完全に移し替えることは原理的に不可能であり、原典と翻訳の関係性に関して、様々な理論の構築や研究が積み重ねられてきている[1,2]。このため古典文献における他言語の影響

を検討する場合には複雑な原典と翻訳の関係性も同時に考慮する必要がある。

これらの翻訳の問題に対して、近年情報処理技術を用いたコーパス言語学や自動翻訳の研究が精力的に進められている[3,4,5,6]。ただしこれらの研究の主たる目的は技術的な文書（科学論文・特許）や Web 上のテキストなどの自動翻訳であり、語の持つ複雑で微細なニュアンス等の要素は重視されていない。このため人文的な古典研究での精密な意味分析には従来の自動翻訳研究は必ずしも有用とは言えない。

本研究ではこれらの背景を踏まえ、まずは最小の意味単位である語彙のレベルに絞って、人文的古典文献研究に利用可能な自動翻訳アルゴリズムを導入し[7]、その結果を用いて他言語テキスト間の概念比較のケーススタディ的分析を行う。

2. 分析の対象と課題

2.1. 分析対象

本研究のケーススタディ分析の対象として聖書を用いる。聖書は旧約聖書と呼ばれるユダヤ教の正典とされる部分が主にヘブライ語で記されているが、新約聖書と呼ばれるキリスト教のみでの正典部分は主にギリシア語で記されている。このためキリスト教神学においてはヘブライ語旧約聖書における神の概念とギリシア語新約聖書における神の概念の対応関係や諸々の神学的命題の相違が問題となってきた。旧約聖書と新約聖書の言語の相違から生じる様々な解釈上の問題に対して、中世ヨーロッパの教会ではヘブライ語でもギリシア語でもないラテン語の翻訳聖書(ウルガタ訳)に権威を与えて、公式の翻訳として用いてきた歴史がある。そのため、古代・中世の神学を学ぶ場合には聖書の原典のみならず、その古代語訳における聖書解釈を把握した上で神学者のテキストに向き合う必要が生じる。

このような非常に複雑な多言語テキストの関係性は聖書学・神学習得の大きな妨げであり、初学者は多くの時間を語学の勉強に充てざるを得ない。類似の状況は他の多数の古典的思想テキストの学習課程においても見られるものであり、テキストの翻訳にまつわる諸問題が情報処理技術の活用によって緩和されればこれらの古典テキストに関わる学問の進歩の助けとなると考えられる。

また、これらの古典テキストの意味解釈において、翻訳テキスト間での意味の相違を計量的に扱うことができるようになれば、より精密で客観的な研究を行うことが可能になると期待される。

本研究では漸近的対応語彙推定法の対象として、ヘブライ語旧約聖書として BHS(Biblia Hebraica Stuttgartensia)[8]を、ギリシア語新約聖書として NA(Nestle Aland)[9]を原典テキストのデータとして用いた。またヘブライ語ギリシア語間の対応関係抽出のための翻訳テキストとして英語の翻訳からは NRSV (New Revised Standard Version)[10]と NAB (New American Bible)[11]を用いた。日本語の翻訳からは「新共同訳聖書」[12]および「口語訳聖書」[13]を用いた。

ヘブライ語・ギリシア語原典の電子テキストと英語翻訳の電子テキストに合わせて、分析に用いる語の活用形の分析データは BibleWorks[14]所収の物を用いた。日本語翻訳の電子テキストは JBible[15]所収の物を用いている。英語テキストの形態素解析には TreeTagger[16]を、日本語電子テキストの辞書作成と形態素解析には MeCab[17]と TextSeer[18]を用いた。

2.2. 分析の方針

本研究では、聖書におけるヘブライ語テキストとギリシア語テキストでの語彙のニュアンスの関係性を、翻訳を仲立ちとすることで抽出し視覚化を行う。

旧約聖書中のヘブライ語の特定の単語が新約聖書中のギリシア語のどの単語と概念的に対応するかは自明ではない。辞書的な意味での対応をとることは可能であるが、抽象的な神学概念やメタファー的な用法に関して、辞書的な意味は有用ではなく、当然のことながら解釈者によって種々に意見が分かれる。

しかしながら現代語の翻訳聖書では、ヘブライ語もギリシア語も現代語に翻訳されており、現代語を挟んで古代語間の対応を抽出することが可能である。これはすなわち翻訳を一つの解釈として用い、翻訳者の理解における古代語間の意味概念の対応関係を分析することにあたる。複数翻訳に対してこのようなプロセスを経ることで翻訳間の神学概念に関する解釈の相違を抽出することが可能であり、また多数の翻訳から平均的な解釈を導くことも考えられる。

3. 漸近的対応語彙推定法

3.1. 漸近的対応語彙推定法の特徴

本研究では、原典と翻訳テキスト間の語彙の対応関係の推定法として、人文的研究に不可欠である語彙の微細なニュアンスの差異を抽出可能な漸近的対応語彙推定法を用いた[7]。

従来の翻訳語彙の対応関係抽出手法においては、自動翻訳で用いることのできる最適な語彙の対応関係を高精度に抽出することが主眼であったため、ある単語が複数の単語に訳し分けられている場合も最も翻訳される確率の高い単語の特定が主たる目的とされてきた。現代語での場合、分析の対象となる原語と翻訳語間のテキスト対は無数に準備可能であるため、大規模なテキストの利用可能性を前提とする場合が多い。また、語彙の対応関係を抽出するデータソースとなるテキスト自体には関心がないため、分析対象テキスト中の単語のどれだけの比率の語彙の対応関係を抽出できたかという再現率に関しても、少量データからの効率的な単語対の抽出という側面以外からは重視されてこなかった。

漸近的対応語彙推定法では、多くの古典テキストの分析でそうであるように、対象テキスト以外には他の大規模テキストコーパスや学習用データの利用が困難なテキストにおいて利用が可能であり、かつ従来の手法に比べて再現率が20%程度向上している。また、最尤推定法等と異なり、一つの単語に対する一つのみの翻訳語を特定するのではなく、複数単語に訳し分けられる場合にもそれらの対応関係を抽出することが可能である。

このため、対象となるコーパス内のテキスト量が必ずしも多くなく、類似単語間の概念の関係性や詳細な単語のニュアンスの分析が必要となる一方で、対象テキスト全体にわたる分析が必要となる古典テキストの分析に適していると考えられる。

3.2. 漸近的対応語彙推定法のアルゴリズム

多くの語彙対応関係推定アルゴリズムにおいては、原文の単語 o と翻訳文の単語 t の間の関係性の指標として相互情報量が用いられる。相互情報量とは、原文で o が出現した場合に対応する翻訳文で t が出現する確率 $P(t|o)$ と、逆に t が出現した場合に対応文で o が出現する確率 $P(o|t)$ の積である。すなわち o と t の相互情報量 Iot は、 $Iot = P(t|o) \times P(o|t)$ と定義できる。基本的には相互情報量が大きいほど、2 単語が翻訳での対応関係にある可能性が高いと推定される。

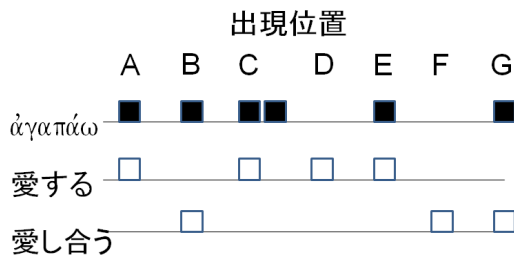


図1 単語の対応する組の関係

図1に例として、ギリシア語の”*ἀγαπάω*”と日本語の「愛する」「愛し合う」での対応関係を示す。この場合”*ἀγαπάω*”はCで二回出現し、CDF以外の箇所では一回ずつ出現している。”*ἀγαπάω*”と「愛する」は箇所ACEで対応しており、これらより相互情報量は、 $3/6 \times 3/4 = 0.375$ となる。同様に”*ἀγαπάω*”と「愛し合う」の相互情報量は約0.22と計算できる。単語の組の抽出での相互情報量の閾値が0.3の場合、”*ἀγαπάω*”と「愛する」が対応する組として抽出され、ACEの3箇所です”*ἀγαπάω*”が合計3回推定されたことになる。閾値が0.2の場合には、ABCEGの5箇所です合計5回”*ἀγαπάω*”の対応単語組が推定され、推定単語の頻度合計は5である。

一般的にはこのように相互情報量で閾値を設け、閾値以上の相互情報量を持つ対応単語の組を抽出する。閾値が高いと対応関係が正しい確率、つまり精度は向上するが、閾値を満たす単語のペアが少なくなり再現率は低下する。基本的に再現率を上げたい場合には相互情報量の小さい単語対までを対応単語として抽出し、精度を上げたい場合には相互情報量の閾値を上げることになる。このような単純な相互情報量のみ

で原典と翻訳間の単語対を抽出した場合、精度は5割強程度と低い[7]。

このため、まず「単語の一対一対応の仮定」をおき精度と再現率の向上を図る。原典と翻訳間の単語対のうちでは、原典で一語の物が翻訳では複数語に翻訳される場合や、その逆の場合がありうる。ただしギリシア語新約聖書と日本語翻訳聖書の場合この比率は1%から5%程度とそれほど大きい値ではない[7]。これより、原典と翻訳の単語が基本的に「一対一対応」していると仮定して単語対の抽出を行う。

単語の一対一対応の仮定に基づく、単語の組の推定のプロセスを図2に例示する。

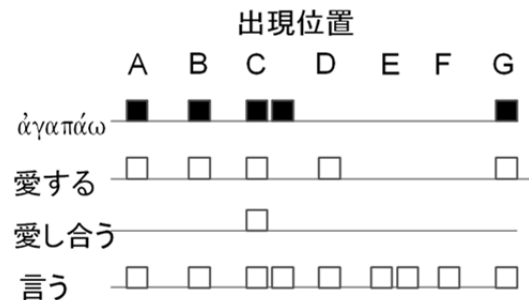


図2 一対一対応仮定での推定1

まず、図2で”*ἀγαπάω*”に対して候補となる翻訳語との相互情報量を計算する。このとき「愛する」とは0.64、「愛し合う」とは0.2、「言う」とは約0.56である。単純に閾値で切ると、0.6以上なら”*ἀγαπάω*”と「愛する」の組のみが、0.5以上ならこれに合わせ”*ἀγαπάω*”と「言う」の組が対応組として抽出される。

これに対し、一対一対応の仮定をおき、閾値の高い物から順に推定していくと、最初に”*ἀγαπάω*”と「愛する」の組が抽出される。次に相互情報量が大いのは「言う」であるが、これは「愛する」と出現位置が重複しており、一対一対応仮定に反するので除外される。そして、次に相互情報量の大い「愛し合う」は出現位置が「愛する」と重複しない(Cの位置は”*ἀγαπάω*”が2回出現しており、2回まで推定が可能である)ため対応組として推定される。相互情報量の閾値が0.2以下ならば結果として”*ἀγαπάω*”の対応組は「愛する」と「愛し合う」と推定される。このように、一対一の仮定をおくことで一つの原文の単語に対し複数の対応組が、共起語を除外する形でより正確に抽出されることが期待される。

「一対一仮定」のみの場合には相互情報量は単語の組を抽出する前の値を用いていたが、正確には被推定単語組の使用箇所を除いた残渣データに対して再度相互情報量の計算を行うべきである。単語組の抽出によって相互情報量の計算における分母の単語数が増えているからである。被推定単語部分を除外して再計算すると、推定された単語の出現頻度の総数が低下し、相対的に相互情報量が上昇

する。再計算によって、相互情報量が閾値を超えた単語の組を再度抽出することにより適切な単語組をより高い閾値でも抽出可能となり率を上昇させられると期待される。

これを先ほどの例で示すと、図 2 での「愛する」と「ἀγαπάω」の対応単語の組が推定された後に、「ἀγαπάω」と「愛する」の対応ですでに使われた箇所を削除して相互情報量を再計算する。図 3 での点線部分が最初の推定で対応組が特定された位置である。これらを除いて相互情報量を再計算すると、「愛する」とは 0, 「愛し合う」とは 1.0, 「言う」とは約 0.1 となり、「愛し合う」と「ἀγαπάω」の組の相互情報量が大幅に上昇していることが分かる。一対一対応の仮定のみでも相互情報量の閾値を下げれば「愛し合う」と「ἀγαπάω」の組が抽出可能であるが、閾値を下げることで妥当な組み合わせでない可能性の高い組も同時に抽出することになり、全体での精度が低下する懸念があった。しかし、相互情報量の再計算によって、相互情報量の閾値を高く保持したまま、複数語義の抽出が可能となる。

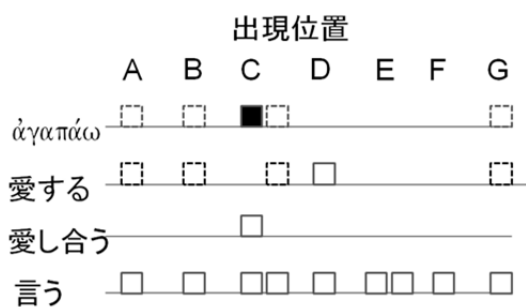


図 3 一対一対応仮定での推定2

これらの手法に組み合わせて、相互情報量の多いものから順に推定を進めていく、すなわち閾値を漸近的に下げていくことで、正しい単語組である確率の高い物から順に推定を進めていくことが可能となる。

これらの手法を組み合わせた結果として、ギリシア語新約聖書と日本語の翻訳聖書からの対応単語組の抽出においては精度 95%以上、再現率 78%以上の単語組の抽出を実現できている[7].

3.3. 抽出された単語のペア数

抽出された言語と翻訳の語彙のペアは、原典テキストと英語翻訳の間では、BHS-NRSV 間で 3501 対、NA-NRSV 間で 2115 対、BHS-NAB 間で 3055 対、NA-NAB 間で 2192 対である。

原典テキストと日本語翻訳の間では、BHS-新共同訳間で 4945 対、NA-新共同訳間で 3951 対、BHS-口語訳間で 4193 対、NA-口語訳間で 4000 対である。

4. 異言語テキスト間概念ネットワーク

4.1. 対応語彙のペアからのネットワーク作成

得られた原語と翻訳語の対応語彙のペアより、翻訳語を仲介として、異なる原語間の対応関係を抽出することが可能である。語彙のペアでの対応関係は数値的な物であり、計量的な分析を施すことが可能である。一方、これらの関係性をネットワーク表現を用いて視覚化することで、複雑なニュアンスの関係性の理解の補助とすることも可能である。

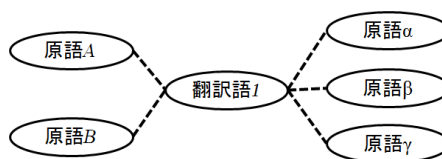


図 4 翻訳語からの距離 1 のネットワーク例

図 4 に、翻訳語と原語の単語の対応関係からネットワークを作る例を示す。図 4 では、「翻訳語 1」に対して、ある原典が記された言語の単語「原語 A」と「原語 B」が対応している。一方で同じ「翻訳語 1」は別の原典が記された別言語の「原語 α」「原語 β」「原語 γ」とも対応関係を持っているとする。この場合図 4 で示したように、これらすべての語との間に線が結ばれたネットワークが構築される。

図 4 では翻訳語からのネットワークを考えたが、翻訳語と対応関係にある原語の各語からも対応関係のネットワークを構築することが可能である。

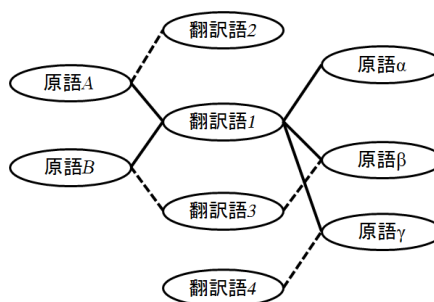


図 5 翻訳語からの距離 2 のネットワーク例

図 5 には、図 4 の例での各原語の語に対して、対応する翻訳語の語を付加したネットワークを示す。この図は最初の翻訳語 1 から見た場合に、距離 2 までの語を含んだネットワークとなる。同様にしてある語から初めて距離 3 や距離 4 のネットワークが必要に応じて作成可能である。

以下本論文では聖書解釈において多様な解釈が展開される語とその解釈に関して、翻訳を仲立ちとして視覚的に表現するケーススタディを紹介する。

4.2. 「神」と「主」の関係性

例として神学で問題となる「主」と「神」（英語では“LORD”と“God”）と原語の単語の対応関係を NRSV の場合を図 6 に、NAB の場合を図 7 に、新共同訳の場合を図 8 に、口語訳の場合を図 9 に示す。

ネットワークのエッジは抽出された単語間のペアである。図の中では“God”および“Lord”から距離 3 までのネットワークを示している。

聖書中では唯一神を「主」と表現したり、「神」と表現したり分けた表現が見られる一方で、「主なる神」と合わせて表現する形も頻繁に見られる。聖書学では唯一神に対する表現の違いの詳細な分析から、旧約聖書における文書仮説などの学説が発達してきた歴史もあり。「主」と「神」（英語では“LORD”と“God”）をどのように理解するかは非常に基本的かつ重要な問題であると言えよう。

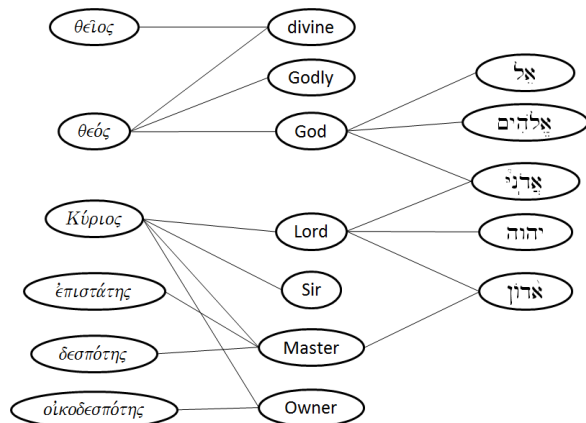


図 6 NRSV での“God”と“Lord”からの距離 3 のネットワーク

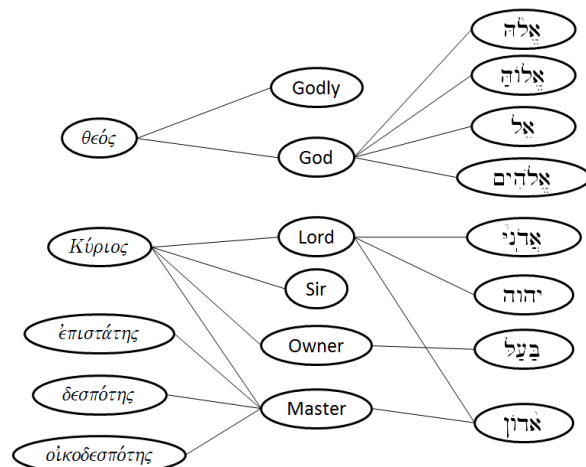


図 7 NAB での“God”と“Lord”からの距離 3 のネットワーク

図 6 と図 7 より、ギリシア語では NRSV と NAB いずれも“God”はθεόςと“Lord”はΚύριοςと対応していることが分かる。一方ヘブライ語では“God”と“Lord”いずれも複数の単語と対応しており、“God”の対応は NRSV と NAB で部分的に異なる。特に、NRSV では אֲדֹנָי を仲立ちとして“God”と“Lord”が接続されており、NRSV での解釈では אֲדֹנָי は“God”と“Lord”の中間概念と分かる。他に共通点としてはいずれの翻訳でも“Master”はΚύριοςと אֲדֹנָי を仲立ちとして“Lord”と接合しており、Κύριος と אֲדֹנָי が“Lord”と“Master”の意味を合わせ持ったニュアンスである点において二つの翻訳の解釈が一致していると分かる。

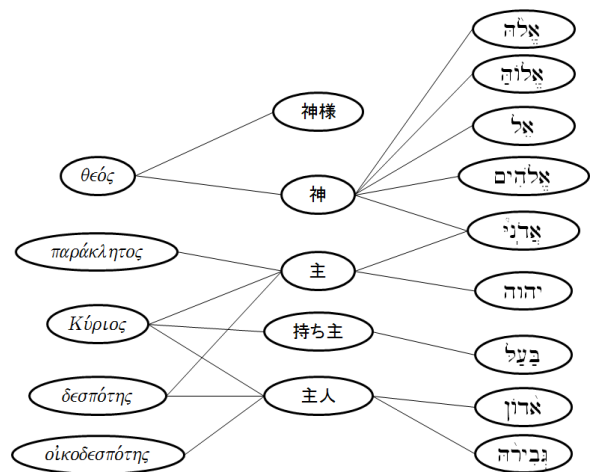


図 8 新共同訳での「神」と「主」からの距離 3 のネットワーク

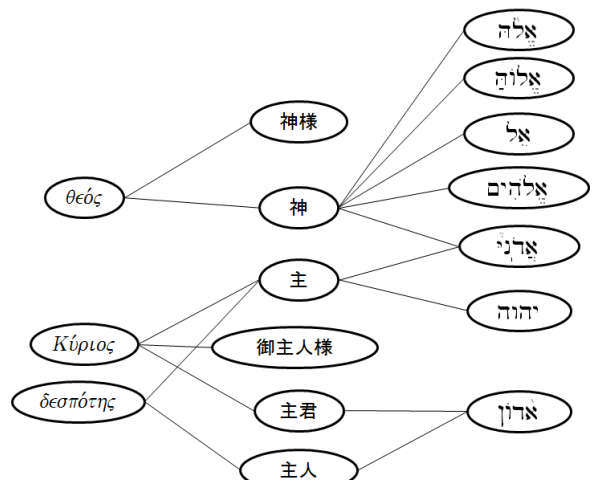


図 9 口語訳での「神」と「主」からの距離 3 のネットワーク

一方日本語訳では、新共同訳口語訳共に NRSV と同じく אֲדֹנָי を仲立ちとして「神」と「主」が接続される構造を持っている。日本語

訳では「神」と「主」に対応するヘブライ語の単語の関係性は全く同一である。ギリシア語に関しては、新共同訳の方がより多くの原語の語が日本語訳の語と結びついている。ただし口語訳での基本的な対応関係は新共同訳にも表れておりこれらの共通性は高いと言えよう。

日本語・英語量翻訳に共通する点としては、多数のヘブライ語の単語が少数の翻訳語と対応関係にあり、複数の翻訳語が少数のギリシア語の単語と対応関係にある傾向が見える。

これはすなわち、ギリシア語に対してはそのニュアンスをより細分化して翻訳で表現しようとするのに対して、ヘブライ語ではいくつかの概念をまとめて翻訳で表現しようとする傾向が見られるということである。

4.2. 「魂」と「命」の関係性

次に多くの宗教で一般的に問題となる「魂」と「命」について考えるためのケーススタディとして「魂」と「命」（英語では“Soul”と“Life”）と原語の単語の対応関係をまず翻訳語からの距離1の場合のネットワークをNRSVを図10に、NABを図11に、新共同訳を図12に、口語訳を図13に示す。

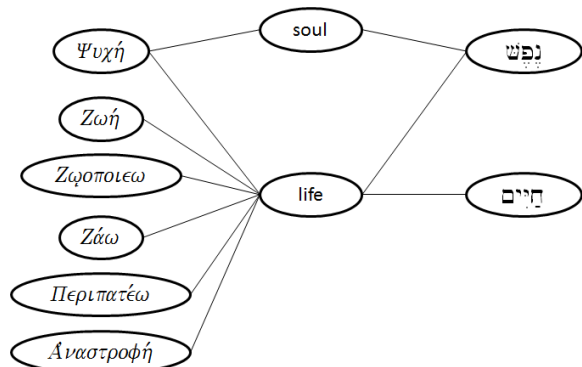


図10 NRSVでの“Soul”と“Life”からの距離1のネットワーク

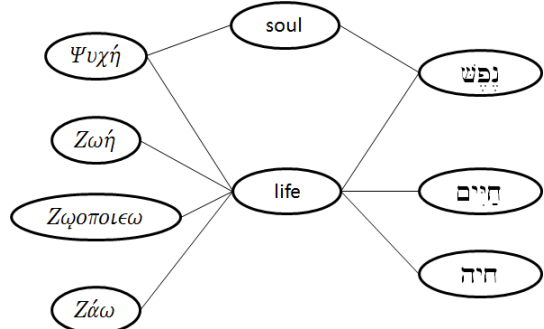


図11 NABでの“Soul”と“Life”からの距離1のネットワーク

図10から13までのすべてのネットワークに共通するのは、「魂」（“Soul”）と「命」（“Life”）の中間概念として、Ψυχήと נפשが

あり、また「命」の訳語として Ζωήと Ζωοποιεωと חייםが対応するという点である。逆に相違点は、各翻訳で対応語の個数が異なる点が挙げられるがその意味する点の詳細は明瞭ではない。そこで次に「魂」と「命」からの距離3のネットワークを構築した結果をそれぞれ図14から図17に示す。

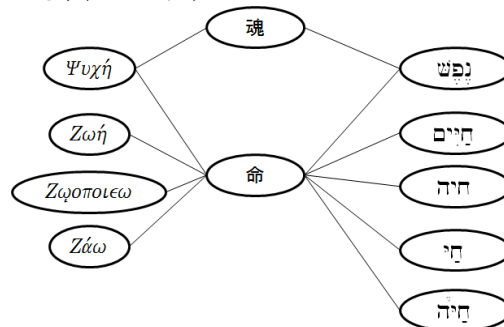


図12 新共同訳での「魂」と「命」からの距離1のネットワーク

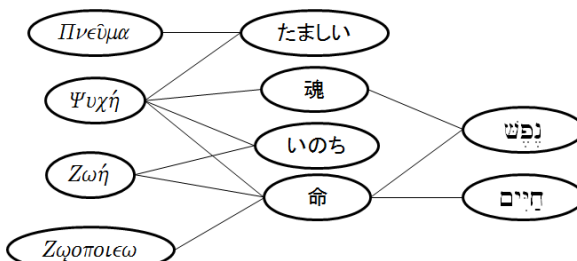


図13 口語訳での「魂」と「命」からの距離1のネットワーク

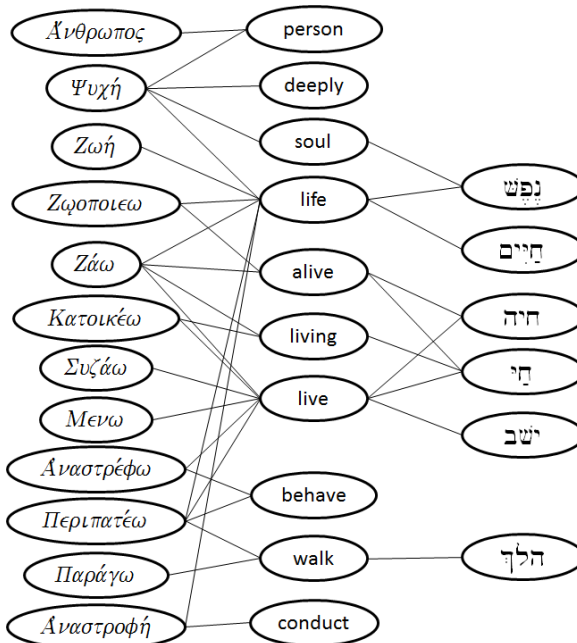


図14 NRSVでの“Soul”と“Life”からの距離3のネットワーク

図 14 より, NRSV における解釈の特徴として, *Ψυχή* が “Person” や “Deeply” と訳されている場合があることが挙げられる. これは「魂」を人間存在の「深い」部分の「人格」そのものとして捉える解釈が背後にあるためと推測できる. また, “Life” が “Behave” や “Conduct” との関連の中で理解されているという傾向も特徴的である. “Walk” を訳語としてあてている *Περιπατέω* や *הלך* はどちらも, 聖書中では単に「歩む」のではなく「どのように生きるか」を示すメタファーとして用いられる場合が多い. これらより, NRSV の解釈では “Life” は生物学的な生命ではなく「どのように生きるか」また「どのように生きるべきか」という問題意識を背後に持った概念を示すため用いられていると考えられる.

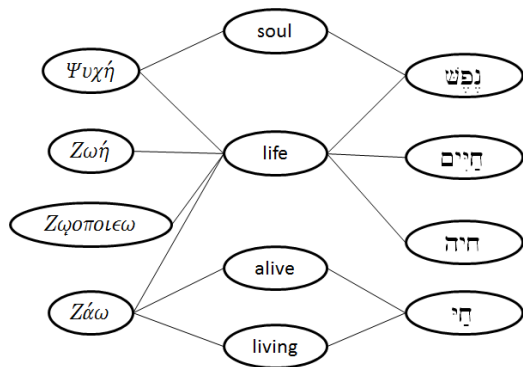


図 15 NAB での “Soul” と “Life” からの距離 3 のネットワーク

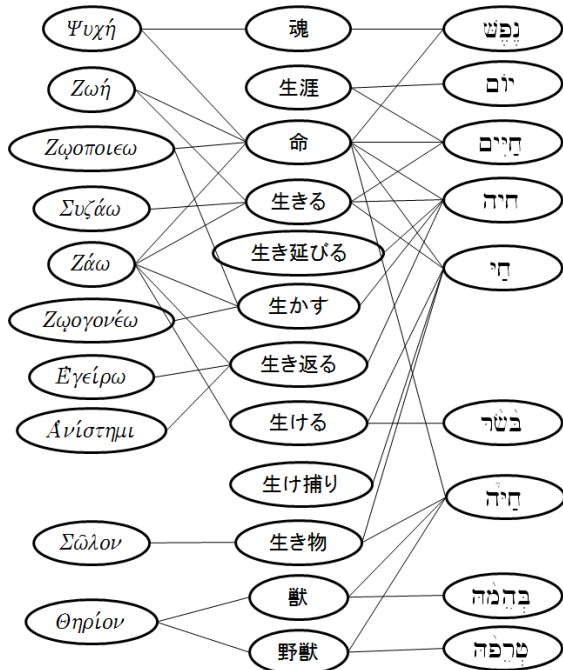


図 16 新共同訳での「魂」と「命」からの距離 3 のネットワーク

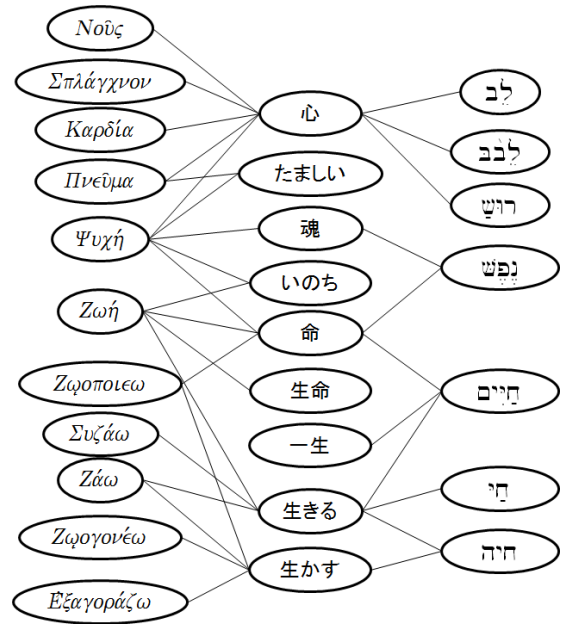


図 17 口語訳での「魂」と「命」からの距離 3 のネットワーク

一方 NAB での距離 3 のネットワーク (図 15) は距離 1 の場合とほとんど変わらない. NAB においては “Soul” と “Life” の概念は他の概念との複雑な関係の中で解釈されず, よりシンプルに単体として理解されていると言えよう. 換言すれば, *Ψυχή* と *נֶפֶשׁ* 以外は個々の原語を基本的に一つの翻訳語に対応させる直訳的方针が取られていると推測される.

日本語訳の新共同訳 (図 16) はギリシア語で「命」の訳語となる *Ζάω* が, 「生き返る」の訳語としても使われる点が他と異なっている. 「命」と「復活」がより密接に関連したものと捉えられているという側面が伺える. また, ヘブライ語で「命」に対応する *חַי* や *חַיָּה* が, 「生き物」や「獣」と重なるニュアンスをもつ語彙として捉えられている点も特徴的である. 「肉体」などと訳されることが多い *בֶּשֶׂר* も周辺に出現している点と合わせると, 身体的・生物学的な生命としての命の側面がより強く表れてきていると言えよう.

口語訳では (図 17), 新共同訳のように「命」に対する踏み込んだ解釈は見られないが, 逆に「魂」に対して, *Ψυχή* や *Πνεῦμα* を通して「心」が接続されている点は特徴的である. 「魂」には「命」につながる身体的な側面がある一方で, 「心」につながる精神的な側面も合わせ持つという解釈が語の対応関係から読み取れる. 特に *Πνεῦμα* は, 「聖」の意味を示す *ἅγιος* と合わせて一般的に「聖霊」とも訳される語であり, 「霊」が「心」であるという, 非常に強く踏み込んだ解釈を読み取ることが可能である.

以上をまとめると、NRSV では“Soul”は人格的側面を含めた概念と捉えられており、口語訳では「命」霊的・精神的側面を含めた概念として捉えられていると言えよう。

また、NRSV では“Life”は「どのように生きるか」という「生き方」のニュアンスを含んでいるのに対し、新共同訳の「命」は生物学的な生命や死からの復活などのニュアンスを含んだ概念であることが読み取れる。

これらに対し「魂」と「命」の概念に関して、NAB は最大公約数的な解釈に基づく翻訳であると言えよう。

5. 結論と今後の課題

本研究では漸近的対応語彙推定法を用い、翻訳を仲立ちとして異言語の古典テキスト間の概念の対応関係をケーススタディ的に抽出し視覚化した。

二つのケーススタディからは、まず「神」と「主」の概念に関してはヘブライ語での多様な語彙を翻訳はより少数の語彙にまとめており、逆にギリシア語の語彙はより多数の翻訳の語彙で説明されるという共通の傾向が見られた。

「神」と「主」の中間的概念として אֱלֹהִים を解釈するか否かに関しては翻訳による差異が現れた。

次に「魂」と「命」の概念に関しては、両者の中間的概念として *Ψυχή* と נֶפֶשׁ があるという点が各翻訳に共通であった。各翻訳の個別的特徴としては、NRSV では“Soul”の人格的な側面が示唆され、口語訳では「命」の霊的・精神的側面が抽出された。また、NRSV では“Life”は「生き方」のニュアンスを含み、新共同訳の「命」は生物学的な生命や死からの復活などのニュアンスを含んでいるとの考察を得た。

このようにネットワーク構造を用いた視覚化によって、大規模テキストを背景とした複雑で微細なニュアンスの特徴をより簡便にかつ客観的データに基づいて把握することが可能である。

ただし、得られた原典と翻訳の語の対応関係性は本来数値的指標であり、今後はその性質を活かした数値的意味解析や構造化手法の確立が課題となると考えられる。語の対応関係を数値的分析手法と組み合わせることで、概念マップやシソーラスの自動構築等に活用できる可能性もある。

また、古典テキスト等のデジタルアーカイブを用いてコンピュータ上で翻訳を閲覧する場合、原語との対応関係、その翻訳に特徴的な解釈等を確認しながら解釈を進めることはテキストのより深い理解に有用と考えられる。そのため、本研究の成果を原典と翻訳のテキストデータと合わせてデジタルアーカイブシステム上に搭載することも今後の課題の一つと言えよう。

謝辞

本研究は科研費「レトリカルデータベースシステムの構築による計量的修辞分析手法の確立」(22700256)の助成を受けた。

参考文献

- [1] 川本皓嗣, 井上健, “翻訳の方法”, 東京大学出版会, 1997.
- [2] 矢崎祐一, “英日翻訳における聖書引用箇所の記事—詩篇 23 篇の場合—”, 翻訳研究への招待, No. 4, pp. 107-122, 2010.
- [3] Peter Brown, Vincent Della Pietra, Stephen Della Pietra, Robert Mercer, “The Mathematics of Statistical Machine Translation: Parameter Estimation”, *Computational Linguistics*, Vol. 19, No. 2, pp.263 - 311, 1993.
- [4] Philipp Koehn, “Europarl: A parallel corpus for evaluation of machine translation”, In *Proceedings of MT Summit*, 2005.
- [5] 宇津呂武仁, 日野浩平, 堀内貴司, 中川聖一, “日英関連報道記事を用いた訳語対応推定”, *自然言語処理*, Vol. 12, No. 5. pp. 43-69, 2005.
- [6] Reinhard Rapp, “Automatic Identification of Word Translations from Unrelated English and German Corpora”, *Proceeding of 37th ACL*, pp.519-526, 1999.
- [7] 村井源, “漸近的対応語彙推定法に基づく翻訳文の解釈的特徴の抽出-日本語翻訳聖書の計量的比較-”, *Vol. 20, No. 3*, pp. 293-310, 2010.
- [8] Deutsche Bible Gesellschaft, “*Biblia Hebraica Stuttgartensia*”, 1997.
- [9] Kurt Aland, et al., “*Nestle-Aland 27th Edition*”, 1993.
- [10] National Council of the Churches of Christ in the USA, “*New Revised Standard Version*”, 1989.
- [11] Confraternity of Christian Doctrine, “*New American Bible*”, 1991.
- [12] 日本聖書協会, “*新共同訳聖書*”, 日本聖書協会, 1987.
- [13] 日本聖書協会, “*口語訳聖書*”, 日本聖書協会, 1955.
- [14] LLC, “*BibleWorks*”, <http://www.bibleworks.com/>, 2012/9/27 参照.
- [15] 日本コンピュータ聖書研究会, “*JBible*”, <http://www.jbible.net/jcbr/>, 2012/9/27 参照.
- [16] Helmut Schmid, “*TreeTagger*”, <http://www.ims.uni-stuttgart.de/projekte/corplex/TreeTagger/>, 2012/9/27 参照.
- [17] 工藤 拓, “*MeCab: Yet Another Part-of-Speech and Morphological Analyzer*”, <http://mecab.sourceforge.net/>, 2012/9/27 参照.
- [18] 川島隆徳, 村井源, “*TextSeer*”, http://www.valdes.titech.ac.jp/~t_kawa/, 2012/9/27 参照.